REMARKS

Claims 37-50 were pending in the subject application. By this Amendment, applicants have hereinabove amended claims 37, 42, 43, 48-50 and added new claim 51.

Applicants maintain that the amendments to the claims are fully supported by the specification as originally filed and do not raise any issue of new matter. Accordingly, applicants respectfully request entry of this Amendment.

Claims Objections

In section 6 of the August 29, 2008 Office Action, the Examiner objected to claims 43 and 48 for informalities. Specifically, the Examiner objected to claim 43 for lacking the recitation "plant" after "selected" and objected to claim 48 for not reciting the full form of "ERECTA."

In response, in order to expedite prosecution but without conceding the correctness of the Examiner's position applicants have amended claims 43 to recite "...wherein the method further comprises propagating the selected plant." Applicants have amended claim 48 to recite, "...transforming a culture of plant cells with the full-form of the ERECTA gene..." Accordingly, applicants respectfully maintain that claims 43 and 48 are not objectional and respectfully request the Examiner reconsider and withdraw this objection.

Claim rejections under 35 U.S.C. §112, Indefinite

In section 7 of the August 29, 2008 Office Action, the Examiner rejected claims 37-50 under 35 U.S.C. §112, second paragraph, as allegedly indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claims 37, 42, and 48

The Examiner alleged that claims 37, 42 and 48 are indefinite for the recitation "...transcribed to form a transcription product which is then expressed..." The Examiner asserted that the transcription product is itself an expression product stating that, "A transcription product is suppose to under translation to produce the protein."

In response, applicants respectfully direct the Examiner's attention to Lubert Stryer, et al. Biochemistry, 2002 5th edition, copy of relevant pages attached as Exhibit A which defines gene expression as "the combined process of the transcription of a gene into mRNA, the processing of that mRNA, and its translation into protein (for protein-encoding genes)." However, to expedite prosecution but without conceding the correctness of the Examiner's position, applicants have herein amended claims 37, 43, and 47 to no longer recite the phrase "which is then expressed."

Claim 48

The Examiner alleged that claim 48 is indefinite for the recitation "gene" which is asserted to be confusing since the limitation "gene" implies that the structure comprises the coding sequence and the associated promoter, terminator and enhancer encoding regions are also a part of the structure (see The Federal Register, Vol. 66, No. 4, Friday, January 5, 2001 at page 1108, left column, Endnote 13). The Examiner indicated that all subsequent recitations of "gene" are also rejected.

In response, applicants respectfully traverse the Examiner's rejection. The Examiner asserted that in the instant case, applicants do not appear to describe such ERECTA gene associated nucleic acid sequences. Applicants respectfully point out that the sequence of the ERECTA gene was known and published prior to the filing date, for example as noted in the present specification at pages 33 and 34 in which alleles of the ERECTA gene in Arabidopsis are identified. Torii *et al.*, (1996) for example provides the nucleotide and deduced amino

acid sequence of the ER gene (see for example Figure 5 and the description of Figure 5 at page 740 of Torii *et al.)* and also in GeneBank/EMBL/DDB/ under accession numbers U47029 and D83257 for the cDNA and genomic DNA sequence, respectively, with the positions of introns and exons. Accordingly, the gene sequence is part of the prior art. Furthermore, the specification describes that an ERECTA gene may be used (at pages 55 to 59 of the specification as filed, paragraphs [0277]-[0296] of the specification as published), and that the presence of a promoter (pages 57-59), and a terminator is clearly contemplated.

Applicant have added new claim 51 reciting all the limitations of claim 48, with the exception that the culture of plant cells are transformed with "...a nucleic acid encoding an ERECTA protein..."

Claim 49 and 50

The Examiner alleged that the recitation of "seeds from a selected plant," in claims 49 and 50 are confusing, asserting that it is unclear whether the seeds comprise the nucleic acid.

In response, to expedite prosecution but without conceding the correctness of the Examiner's position, applicants have herein amended claims 49 and 50, to recite, "...wherein said seeds comprise the nucleic acid."

Accordingly, applicants respectfully request that the Examiner reconsider and withdraw these anticipation rejections.

Rejections under 35 U.S.C. §112, Written Description

In section 8 of the August 29, 2008 Office Action, the Examiner rejected claim 48 under 35 U.S.C. §112, first paragraph, as allegedly not complying with the written description requirement. The Examiner alleged that the claim contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventors, at the time the application was filed, had possession of the claimed invention. The Examiner's detailed reasons are at forth on page 5-7 of the August 29, 2008 Office Action.

Applicants' Reply

In response, applicants respectfully traverse the Examiner's rejection. The Examiner asserts that claim 48 is not sufficiently described, as the specification does not describe the structure of ERECTA genes isolated from diverse sources and genetic backgrounds. Applicants respectfully note that the Examiner has not acknowledged that Examples 12, 14 and 15 set out in detail methods which were used to identify orthologs of *Arabidopsis* ERECTA in sorghum, wheat and maize, and in Example 13 ERECTA homologs in *Arabidopsis thaliana*. These homologs and orthologs are also set out in Figures 12 to 15 of the present specification.

As discussed above, ERECTA was known prior to the filing date of the present application. In studies relating to the characteristics and phylogeny of Receptor-Like Kinases in *Arabidopsis*, Shiu and Bleeker (2001), attached hereto as **Exhibit B**, found that there were about 200 Leucine-Rich-Repeat Receptor like Kinases (LRR-RLK) which were easily recognisable by the simple sequence analysis for the presence of "repeats" in the receptor domain. Within the LRR-RLK, further division into

subfamilies were made on the basis of other structural/sequence features, and Shiu and Beeker provides the precise characteristics necessary to identify respective LRR-RLKs. ERECTA is a member of the LRR-RLK group, belonging to subclass LRR XIII set out in Figure 2 at page 10766 of Shiu and Beeker (2001). Figure 4 of this paper identifies only 7 members of this LRR XIII subclass in *Arabidopsis*.

Using methods described in the present specification, for instance as exemplified by the applicants in Examples 12 to 15, a person of skill in the art at the filing date would readily be able to identify and characterise ERECTAs from amongst the family of Leucine-Rich-Repeat Receptor-Like Kinases.

The Examiner asserts that Shpak *et al.* (2004) suggests that ERECTA genes are involved in diverse cellular processes, and thus the present broad claim encompasses structures whose function is unrelated to the ERECTA polypeptide of SEQ ID NO:2.

In response it is noted that the Examiner appears to have assumed that a diversity of function of ERECTA is based in a diversity of structure. This assumption is not supported by fact. For example, the Examiner has asserted in paragraph 9 of the Office Action that Japanese patent publication No. JP 09056382A ("Mitsukawa") discloses the use of a polypeptide of SEQ ID NO:2. Mitsukawa discloses that a polynucleotide encoding a polypeptide of SEQ ID NO: 2 is associated with the control of elongation of plant stems. In the present specification, however, the polypeptide of SEQ ID NO: 2 is demonstrated to be associated with the distinct and unrelated function of transpirational efficiency.

It is respectfully submitted that the same ERECTA may therefore be involved in a diversity of functions. This diversity of function may, for example, be a result of the region of the plant genome into which the ERECTA is introduced, the copy number of

the ERECTA which is expressed, and the degree to which the plant's endogenous ERECTA is deficient or sub optimal. It may also be related to the cell and tissue specificity of the ERECTA protein expression. Thus the present specification reveals that ERECTA regulates stomatal density (see for example Figure 2a) , and hence is involved in a pathway associated with stomatal patterning, which is quite distinct from a pathway regulating elongation of cells in the reproductive stems. The claims of the present specification, but not the prior art, require the selection of plants having enhanced transpirational efficiency. Suitable methods for identifying the level of transpirational efficiency of a plant are described in detail in the specification.

Accordingly, it is respectfully submitted that the specification fully describes claim 48 and applicants request that the Examiner reconsider and withdraw the rejection for lack of written description.

Rejections under 35 U.S.C. §§102 (b) and 103(a)

In section 9 of the August 29, 2008 Office Action, the Examiner rejected claims 37-50 under 35 U.S.C. §102(b) as allegedly anticipated by or, in the alternative, under 35 U.S.C. §103(a) as allegedly obvious over Mitsukawa et al. (Japanese Patent Publication No. JP 09056382 A, published March 4, 1997) and evidenced by Masle et al. (Nature, 436:866-870, 2005).

Applicants' Response

Rejections under 35 U.S.C. §102 (b)

In response, applicants respectfully traverse the Examiner's rejection. The Examiner asserts that Mitsukawa discloses a method of producing a transgenic plant which would inherently possess the properties possessed by the presently claimed

invention.

The Examiner has misapplied the law of inherent anticipation, and consequently has made an improper Rejections under 3 5 U.S.C. §102.

As cited in M.P.E.P. §2112 with regard to inherent anticipation, "[t]he fact that a certain result or characteristic may occur or be present in the prior art is not sufficient to establish the inherency of that result or characteristic. In re Rijckaert, 9 F.3d 1531, 1534, 28 USPQ2d 1955, 1957 (Fed. Cir. 1993)". More specifically, "[t]o establish inherency, the extrinsic evidence 'must make clear that the missing descriptive matter is necessarily present in the thing described in the reference, and that it would be so recognized by persons of ordinary skill. Inherency, however, may not be established by probabilities or possibilities. The mere fact that a certain thing may result from a given set of circumstances is not sufficient.' In re Robertson, 169 F.3d 743, 745, 49 USPQ2d 1949, 1950-51 (Fed. Cir. 1999)" (M.P.E.P. §2112) (emphasis added). Accordingly, the Examiner's unsupported statement that "such a property [enhanced transpiration efficiency] would be inherent to the method of expressing the protein (accession no. AAW13408) in Mitsukawa et al. transgenic plant" is insufficient basis for an anticipation rejection based on inherency. The Examiner has failed to cite any evidence that enhanced transpiration efficiency would necessarily result.

Applicants respectfully submit a verified English translation of Mitsukawa, together with a copy of the complete Japanese document which was translated for the Examiner's convenience, and as a substitute for the computer translation of this citation which the Examiner has used previously are attached hereto as **Exhibits C, D and E** respectively.

Applicants respectfully submit that one of ordinary skill in the art at time of filing would not necessarily arrive at a plant

with enhanced transpirational efficiency when following the method provided by Mitsukawa.

As submitted in previous responses, Mitsukawa does not disclose the step of selecting a plant with enhanced transpirational efficiency as required by the claims. Instead, Mitsukawa discloses a different step of identifying a plant with enhanced stem length. Applicants submit that there is no correlation between stem length and transpirational efficiency in plants with a given ERECTA.

The applicants provide below evidence supporting the lack of correlation.

Firstly, the applicants demonstrate that polymorphisms in ERECTA sequence which lead to significant changes in morphogenesis (as described in the Mitsukawa) do not necessarily lead to changes in transpiration efficiency. The graph which is attached hereto as **Exhibit F** shows the relationship obtained by the applicants between transpiration efficiency and stem length among a range of *Arabidopsis* accessions which are polymorphic in ERECTA. The Y axis plots the plant's ^{13}C isotope discrimination, as a measure of transpirational efficiency (with decreasing ^{13}C discrimination associated with increased transpirational efficiency either due to more closed stomata or to increased photosynthetic capacity or to both as is the case for ERECTA). The X axis plots the height of the mature plant as measured by the height of the inflorescence. Plots of individual accessions are set out as individual points. This graph shows that for any given plant stem height, a range of transpirational efficiencies may be obtained, with no correlation between stem height and ^{13}C discrimination. No statistically significant association between ^{13}C discrimination and plant height was found over the whole range. Thus, in the light of this data, it is submitted that for any given ERECTA allele there is no correlation between the plant

height which is obtained and the transpirational efficiency of that plant.

Secondly, the applicants demonstrate that while the degree of overexpression of ERECTA tends to correlate with increased transpirational efficiency, it does not correlate with plant height. Two graphs are provided attached hereto as **Exhibit G.** These graphs plot the relationship between the degree of ERECTA expression in ERECTA overexpressing plant lines (fold change in expression compared to ERECTA expression in Columbia allele (normalised to a control gene) on X-axis) and either ^{13}C discrimination (as a measure of transpirational efficiency, top graph) or plant height at maturity (bottom graph) on the respective Y axes.

Although there is a trend for plants with increasing ERECTA expression to have enhanced transpirational efficiency beyond wild type levels, the bottom graph shows that apart from null expression lines there was no clear correlation between the level of ERECTA over expression and plant height. Accordingly, the selection of a plant with increased height does not necessarily result in the selection of a plant with enhanced transpirational efficiency.

Accordingly it is respectfully submitted that the disclosure of a process for making plants with enhanced stem length in Mitsukawa does not inherently disclose the methods claimed in the present application.

Rejections under 35 U.S.C. §102(b)

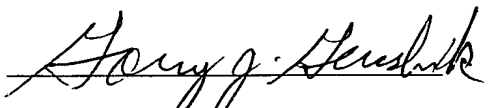
Further, in the light of the above, it is respectfully submitted that Mitsukawa also does not render the claims obvious. The Examiner stated that, "It would have been obvious to one of ordinary skill in the art to select for transgenic plant with increased transpiration efficiency (inherently associated property of polynucleotidesequence disclosed in the reference)

because selection of transgenic plant with a phenotype would have been the ultimate useful goal without any surprises or unexpected results." Applicants are unclear as to what is being referred to as "phenotype." The Examiner does not segregate the different phenotypes at issue here and simply blurs them in a genetic "phenotype," thereby ignoring what applicants are actually claiming. As such the rejection does not deal with claimed invention and is fatally defective. If stem length is the phenotype one would aim to achieve, one skilled in the art would not necessarily obtain the claimed invention, i.e. a plant with enhanced transpirational efficiency. Mitsukawa teaches the selection of plants with increased height, which as set out above, does not necessarily correlate with transpirational efficiency. The Examiner has not identified any evidence which teaches that ERECTA is associated with transpirational efficiency. Applicants have on the other hand presented evidence showing that promotional efficiency does not correlate with stem length.

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicants' undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee, other than the enclosed $1,110.00 fee for a three-month extension of time is deemed necessary in connection with the filing of this Amendment. However, if any additional fee is required authorization is hereby given to charge the amount of any such fee to Deposit Account No. 03-3125.

Respectfully submitted,



John P. White
Registration No. 28,678
Gary J. Gershik
Registration No. 39,992
Attorneys for Applicants
Cooper & Dunham LLP (Customer #23432)
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

I hereby certify that this correspondence is being deposited on this date with the U.S. Postal Service with sufficient postage as first class mail in an envelope addressed to:

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450 Alexandria,
VA 22313-1450

3/2/09
Date

Gary J. Gershik
Reg. No. 39,992

EXHIBIT A

W. H. FREEMAN AND COMPANY
BIOCHEMISTRY
FIFTH EDITION
Jeremy M. Berg John L. Tymoczko Lubert Stryer

31. The Control of Gene Expression

Bacteria are highly versatile and responsive organisms: the rate of synthesis of some proteins in bacteria may vary more than a 1000-fold in response to the supply of nutrients or to environmental challenges. Cells of multicellular organisms also respond to varying conditions. Such cells exposed to hormones and growth factors will change substantially in shape, growth rate, and other characteristics. Moreover, many different *cell types* are present in multicellular organisms. For example, cells from muscle and nerve tissue show strikingly different morphologies and other properties, yet they contain exactly the same DNA. These diverse properties are the result of differences in gene expression.

Gene expression is the combined process of the transcription of a gene into mRNA, the processing of that mRNA, and its translation into protein (for protein-encoding genes). A comparison of the gene-expression patterns of cells from the pancreas, which secretes digestive enzymes, and the liver, the site of lipid transport and energy transduction, reveals marked differences in the genes that are highly expressed (Table 31.1), a difference consistent with the physiological roles of these tissues.

How is gene expression controlled? Gene activity is controlled first and foremost at the level of transcription. Much of this control is achieved through the interplay between proteins that bind to specific DNA sequences and their DNA-binding sites. In this chapter, we shall see how signals from the environment of a cell can alter this interplay to induce changes in gene expression. We first consider gene-regulation mechanisms in prokaryotes and particularly in *E. coli*, because these processes have been extensively investigated in this organism. We then turn to eukaryotic gene regulation. In the chapter's final section, we explore mechanisms for regulating gene expression past the level of transcription. ↑ TOP

Applicant: Josette Masle et al.
Serial No.: 10/519,135
Filed: August 15, 2005
Exhibit A

Tables

Table 31.1. Highly expressed protein-encoding genes of...

Search

GO

This book All books

PubMed

EXHIBIT B

Receptor-like kinases from *Arabidopsis* form a monophyletic gene family related to animal receptor kinases

Shin-Han Shiu and Anthony B. Bleecker*

Department of Botany and Laboratory of Genetics, University of Wisconsin, Madison, WI 53706

Edited by Elliot M. Meyerowitz, California Institute of Technology, Pasadena, CA, and approved July 6, 2001 (received for review March 22, 2001)

Plant receptor-like kinases (RLKs) are proteins with a predicted signal sequence, single transmembrane region, and cytoplasmic kinase domain. Receptor-like kinases belong to a large gene family with at least 610 members that represent nearly 2.5% of *Arabidopsis* protein coding genes. We have categorized members of this family into subfamilies based on both the identity of the extracellular domains and the phylogenetic relationships between the kinase domains of subfamily members. Surprisingly, this structurally defined group of genes is monophyletic with respect to kinase domains when compared with the other eukaryotic kinase families. In an extended analysis, animal receptor kinases, Raf kinases, plant RLKs, and animal receptor tyrosine kinases form a well supported group sharing a common origin within the superfamily of serine/threonine/tyrosine kinases. Among animal kinase sequences, *Drosophila* Pelle and related cytoplasmic kinases fall within the plant RLK clade, which we now define as the RLK/Pelle family. A survey of expressed sequence tag records for land plants reveals that mosses, ferns, conifers, and flowering plants have similar percentages of expressed sequence tags representing RLK/Pelle homologs, suggesting that the size of this gene family may have been close to the present-day level before the diversification of land plant lineages. The distribution pattern of four RLK subfamilies on *Arabidopsis* chromosomes indicates that the expansion of this gene family is partly a consequence of duplication and reshuffling of the *Arabidopsis* genome and of the generation of tandem repeats.

The ability to perceive and process information from chemical signals via cell surface receptors is a basic property of all living systems. In animals, the family of receptor tyrosine kinases (RTKs) mediates many signaling events at the cell surface (1, 2). This class of receptors is defined structurally by the presence of a ligand-binding extracellular domain, a single membrane-spanning domain, and a cytoplasmic tyrosine kinase domain. In plants, receptor-like kinases (RLKs) are a class of transmembrane kinases similar in basic structure to the RTKs (3). In *Arabidopsis* alone, it has been reported that there are more than 300 RLKs (4, 5). In the limited cases where a functional role has been identified for plant RLKs, they have been implicated in a diverse range of signaling processes, such as brassinosteroid signaling via BRI1 (6), meristem development controlled by CLV1 (7), perception of flagellin by FLS2 (8), control of leaf development by Crinkly4 (9), regulation of abscission by HAESA (10), self-incompatibility controlled by SRKs (11), and bacterial resistance mediated by Xa21 (12). Putative ligands for SRK (13, 14), CLV1 (15, 16), BRI1 (17), and FLS2 (18) have recently been identified. Proteins interacting with the kinase domains of RLKs *in vitro* have also been found (19–21).

Plant RLKs can be distinguished from animal RTKs by the finding that all RLKs examined to date show serine/threonine kinase specificity, whereas animal receptor kinases, with the exception of transforming growth factor-β (TGF-β) receptors, are tyrosine kinases. In addition, the extracellular domains of RLKs are distinct from most ligand-binding domains of RTKs identified so far (1, 2). These differences raise the question of the specific evolutionary relationship between the RTKs and RLKs within the recognized superfamily of related eukaryotic serine/threonine/tyrosine protein kinases (ePKs). An earlier phylogenetic analysis (22), using the six RLK sequences available at the time, indicated a close relationship between plant sequences and animal RTKs, although RLKs were placed in the "other kinase" category. A more recent analysis using only plant sequences led to the conclusion that the 18 RLKs sampled seemed to form a separate family among the various eukaryotic kinases (23). The recent completion of the *Arabidopsis* genome sequence (5) provides an opportunity for a more comprehensive analysis of the relationships between these classes of receptor kinases.

To understand the evolution of the RLK family and its relationship with other kinase families and provide a framework to facilitate the prediction of RLK function, we set out to conduct a genome-wide survey of RLK-related sequences in *Arabidopsis*. Through a phylogenetic analysis of the conserved kinase domains, we sought to determine (*i*) whether RLKs belong to a monophyletic group when compared with other ePKs and (*ii*) how the RLKs are related to animal receptor kinases. To investigate the relationship between the evolution of land plants and the expansion of the RLK family, we performed a survey of expressed sequence tags (ESTs) for a variety of organisms. Finally, we looked into the chromosomal distribution of four RLK subfamilies to investigate the potential mechanisms contributing to the expansion of this gene family in *Arabidopsis*.

Materials and Methods

Sequence Selection. *RLKs.* All published plant RLK sequences were retrieved, and their kinase domain sequences were used to conduct batch BLAST analysis (24) for related sequences in Viridiplantae, with an E value cutoff of 1×10^{-10}. The cutoff was chosen based on multiple phylogenetic analyses using data sets generated from cutoff E values of 1×10^{-20}, 1×10^{-10}, and 1. All known RLKs were recovered at 1×10^{-20}; therefore, a more relaxed criterion, 1×10^{-10}, was used to retrieve all potentially related genes. The search results were merged, and redundant sequences were deleted. As of February 2001, more than 900 nonredundant candidates of plant RLKs or related kinases were present in GenBank, and they were used for subsequent phylogenetic analysis. For a complete list of genes in the RLK/Pelle gene family, see supporting information, which is published on the PNAS web site, www.pnas.org. The gene name or accession numbers for RLKs shown in the manuscript are as follows: ARK2 (AAB33486), At2g15300 (AAD26903), At2g19130 (AAD12030), At2g24370 (AAD18110), At2g33580

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: RLK, receptor-like kinase; RLCK, receptor-like cytoplasmic kinases; ePK, eukaryotic protein kinase; RTK, receptor tyrosine kinase; RSK, receptor serine/threonine kinase; APH(3')III, aminoglycoside kinase III; EST, expressed sequence tag; TGF-β, transforming growth factor-β; LRR, leucine-rich repeat.

*To whom reprint requests should be addressed. E-mail: bleecker@facstaff.wisc.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

(AAB80675), At2g45340 (AAB82629), At4g11480 (CAB82153), At4g26180 (CAA18124), At4g39110 (CAB43626), BRI1 (AAC49810), CLV1 (AAF26772), ERECTA (AAC49302), HAESA (CAB79651), PR5K (AAC49208), PERK (AAD43169), RPK1 (AAD11518), RKF1 (AAC50043), RKF3 (AAC50045), RKL1 (AAC95351), TMK1 (JQ1674), WAK1 (CAA08794), F1P2.130 (CAB61984), F4I18.11 (T02456), F13F21.28 (AAD43169), F15A17.170 (CAB86081), F17J16.160 (CAB86939), F18L15.120 (CAB62031), F23E13.70 (CAA18124), F23M19.11 (AAD39611), F27K19.130 (CAB80791), MLD14.2 (BBA99679), and T20L15.220 (CAB82765).

Representatives of the eukaryotic protein kinase (ePK) superfamily. Based on Hanks and Hunter (22), plant and animal sequences from each ePK family were chosen. Plant kinases that seemed to be unique to plants were also included in this study (23). Their accession numbers are as follows. *Arabidopsis* sequences: CDC2a (AAB23643), CPK7 (AAB03247), CKI1 (CAA55395), CKA1 (BAA01090), AME2 (BAA08215), MKK3 (BAA28829), MEKK1 (BAA09057), NAK (AAA18853), NPH1 (AAC01753), PVPK-like PK5 (BAA01715), CTR1 (AAA32779), MRK1 (BAA22079), S6K-like PK1 (AAA21142), GSK3β (CAA64408), GSK3ι (CAA68027), SnRK2-like PROKINa (AAA32845), and Tousled (AAA32874); human sequences: CaMK1 (NP003647), CDK3 (NP001249), CK1α1 (NP001883), CK2α (CAB65624), GRK6 (P43250), RK (Q15835), Hunk (NP055401), CLK1 (P49759), MAPK10 (P53779), MAPKK1 (Q02750), MAPKKK1 (Q13233), cAPK (P17612), Raf1 (TVHUF6), c-SRC (P12931), TLK1 (NP036422), and TTK (A42861).

Animal receptor kinases. One representative human receptor tyrosine kinase sequence was selected from each RTK subfamily (1, 2) as follows: AXL (NP001690), DDR (Q08345), EGFR (P00533), EPH (P21709), FGFR2 (P21802), HGFR (P08581), IR (NP000199), KLG-like PTK7 (AAC50484), LTK (P29376), MuSK (AAB63044), PDGFRβ (PFHUGB), RET (S05582), RYK (I37560), TIE (P35590), TRKα (BAA34355), and VEGFR (P17948). Human TGF-β receptors (TGFβR I, P36897; TGFβR II, P37173) were chosen as animal representatives of receptor serine/threonine kinases.

Sequence Annotation, Alignment, and Phylogenetic Analysis. *Delineation of structural domains.* Structural domains of all sequences were annotated according to SMART (25) and Pfam (26) databases. The receptor-like kinase configuration was determined by the presence of putative signal sequences and extracellular domains. Sequences without signal sequences, transmembrane regions, or putative extracellular domains were also included in the analysis. The kinase domain sequences delineated initially according to sequence prediction databases were modified to include missing or exclude excessive flanking sequences according to the subdomain signature of eukaryotic kinases (22).

Alignment of sequences. The sizes of the kinase domains range from 250 to 300 aa. These sequences were compiled and aligned by using CLUSTALX (27). The weighing matrices used were BLOSUM62 or PAM250 with the penalty of gap opening 10 and gap extension 0.2. The alignments generated by these two scoring tables are similar to each other and were manually adjusted according to the subdomain signatures of eukaryotic kinases (22). The alignment for all 610 RLK family members is provided as supporting information.

Optimality criterion and PAUP program parameters. The aligned sequences were analyzed with PAUP (29) based on the Neighbor–Joining method (28), minimal evolution, and maximum parsimony criteria. To obtain the optimal trees, bootstrap analyses were conducted with 100 replicates using the heuristic search option. Two character-weighing schemes used were (*i*) all characters of equal weight and (*ii*) consider the number of nucleotide changes required to change from one amino acid to the other. All other parameters for PAUP were the default values. Because of the difficulty in aligning kinase subdomain X, two character sets were defined with or without kinase subdomain X sequences.

Tree rooting and display. Aminoglycoside kinase (APH(3')III) from *Staphylococcus* (P00554) (30) and the *Arabidopsis* homolog of RIO1 family kinases (S61006) (31) were used as outgroups in this study. In all analyses, the rooting based on either sequence gave the same results. The numbers associated with each branch represent the bootstrap support, and branches with less than 50% support are collapsed.

Table 1. The proportion of EST records representing RLK/Pelle homologs in various organisms

Organism	Total EST*	RLK homologs	%EST
Porphyra yezoensis	10,185	0	0
C. elegans	109,095	0	0
D. melanogaster	95,211	3	0.003
Chlamydomonas reinhardtii	55,860	0	0
Marchantia polymorpha	1,307	1	0.077
Mosses	9,159	19	0.207
Ceratopteris richardii	2,838	7	0.247
Pinus taeda	21,797	100	0.459
Arabidopsis thaliana	112,467	620	0.551
Glycine max	122,843	704	0.573
Lotus japonicus	26,844	135	0.503
Lycopersicon esculentum	87,680	526	0.6
Oryza sativa	62,390	185	0.297
Triticum aestivum	44,132	178	0.403
Zea mays	73,965	135	0.183

*The searches were conducted based on EST available from GenBank as of Dec. 15, 2000.

Identification of Sequences Representing RLK Homologs. *Genomic sequences.* The kinase domain protein sequences of CLV1 and NAK were used to conduct BLAST searches against the genome sequences of *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster*, and human. The genomic sequence hits with an E value smaller than 1×10^{-10} were included for further analysis. Phylogenetic trees were constructed with the candidate sequences and the eukaryotic protein kinase representatives. Sequences that fell into the same clade as RLKs and had more than 50% bootstrap support were regarded as RLK homologs. The sequences shown in this analysis are *Caenorhabditis* Pelle-like sequence (CePelle, T23534), *Drosophila* Pelle (DmPelle, Q05652), and human IRAK1 (NP001560).

EST sequences. CLV1 and DmPelle kinase domain sequences were used to conduct BLAST searches against the EST records of organisms listed in Table 1. All EST sequences with E values smaller than 1.0 were retrieved for further analysis. The sequences with E values smaller than 1×10^{-50} were regarded as RLK homologs. The rest of the sequences longer than 300 nucleotides were submitted for batch BLASTX searches against *Arabidopsis* polypeptide records in GenBank. These sequences were regarded as RLK homologs if the top five matches of the BLAST outputs were RLK family kinases.

Results

The Diversity of RLKs in the *Arabidopsis* Genome. As the *Arabidopsis* genome sequencing effort approached completion, we conducted a genome-wide survey of the RLK gene family to gain more understanding of its size and complexity. The kinase domains of 22 different plant RLKs with various extracellular domains were used to search for similar sequences in GenBank polypeptide records of Viridiplantae, including all land plants

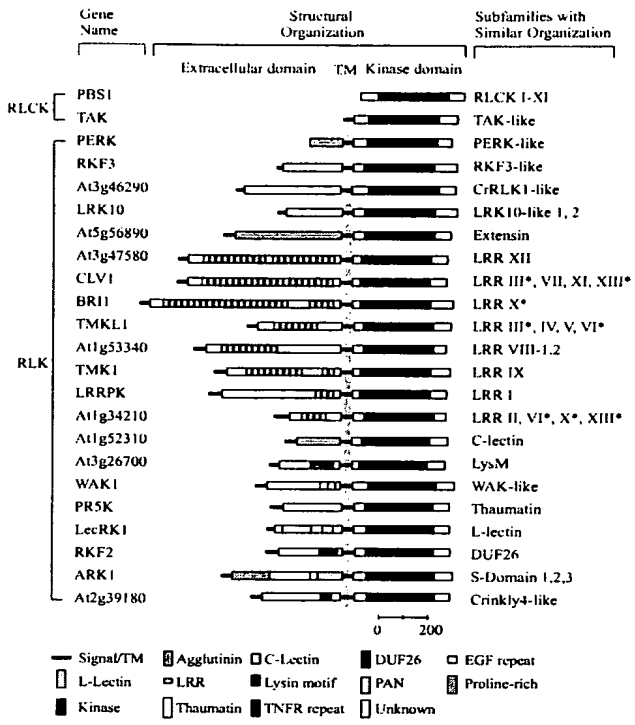


Fig. 1. Domain organization of representative RLKs and RLK-subfamily affiliations. Based on the presence or absence of extracellular domains, members of this gene family are categorized as RLKs or RLCKs. The gray line indicates the position of the membrane-spanning domain. The signal peptides are presumably absent in mature proteins but are displayed to demonstrate their presence in the RLKs. Locus names or MAtDB gene names are provided for the RLK representatives. Domain names are given according to SMART and Pfam databases (25, 26). Subfamilies are assigned based on kinase phylogeny (see supporting information for subfamily assignments for all members of the *Arabidopsis* RLK/Pelle family) and are shown according to the domain organization of the majority of members in a given subfamily. Subfamilies with >30% of members in more than one major extracellular domain category are designated with asterisk. DUF, domain of unknown function; EGF, epidermal growth factor; C lectin, C-type lectin; L-lectin, legume lectin; PAN, plasminogen/apple/nematode protein domain; TM, transmembrane region; TNFR, tumor necrosis factor receptor.

and algae. With the cutoff E value of 1×10^{-10}, more than 900 nonredundant sequences were retrieved. The most recent survey of the completed *Arabidopsis* genome revealed 620 sequences related to RLKs. Ten of these sequences showed greatest sequence similarity to the Raf kinase family. For the remaining 610 *Arabidopsis* sequences, 193 did not have an obvious receptor configuration as determined by the absence of putative signal sequences and/or transmembrane regions (see supporting information). The other 417 genes with receptor configurations can be classified into more than 21 structural classes by their extracellular domains with examples shown in Fig. 1. The sizes of these classes varied greatly. The leucine-rich repeat (LRR) containing RLKs represented the largest group in *Arabidopsis* with 216 genes.

To determine whether RLKs with similar extracellular domains also have similar kinase domains, the polypeptide sequences of the kinase domains of all 620 *Arabidopsis* genes were aligned, and a phylogenetic tree was generated with the Neighbor–Joining method (28) using APH(3')III as outgroup (see supporting information for the complete alignment). APH(3')III is a bacterial gene that is thought to be a distant relative of ePK (30). The phylogeny of *Arabidopsis* kinase domain sequences revealed an interesting pattern where the sequences clearly fell into distinct clades (see supporting information for the phylogenetic tree). We have tentatively assigned these natural groups into 44 different RLK subfamilies based on the kinase domain phylogeny (see supporting information for the subfamily assignment). A noteworthy feature of the pattern obtained is that the members within each of the RLK subfamilies tend to have similar extracellular domains, indicating that a single domain-shuffling event may have led to the founding of each of the various RLK subfamilies. For example, the diverse LRR-containing RLKs fell into more than 13 subfamilies based on kinase-domain phylogeny. With few exceptions, the pattern obtained is consistent with the grouping based on the structural arrangement of LRRs and the organization of introns in the extracellular domains of the individual RLKs (data not shown). Phylogenetic trees were also generated using minimum evolution and maximum parsimony criterions. The results were similar to phylogeny generated with the Neighbor–Joining method (data not shown).

The Relationship Between RLKs and Other Families of Protein Kinases from *Arabidopsis*. Despite the similar domain organization between different plant RLKs, the phylogenetic relationships among members of this family have not been thoroughly studied. Members of the RLK family could have arisen independently multiple times from distinct families of ePKs. Alternatively, they could have originated from a single ePK family and have a monophyletic origin. To address this question, we conducted a phylogenetic analysis by using the kinase domain amino acid sequences of representative RLK sequences from each RLK subfamily and representatives from different ePK families found in *Arabidopsis*.

In the phylogeny based on minimal evolution criterion, all RLK representative sequences from *Arabidopsis* formed a well supported clade, indicating that RLKs have a monophyletic origin within the superfamily of plant kinases (Fig. 2). In addition to RLK sequences, this monophyletic group also included kinases with no apparent signal sequence or transmembrane domain, and they were collectively named receptor-like cytoplasmic kinases (RLCKs, Fig. 1). Some of these kinases formed subfamilies distinct from other RLKs, whereas others were embedded within several different RLK subfamilies. To determine whether the monophyletic grouping of the RLK family represented a bias because of the exclusive use of *Arabidopsis* sequences, an extended analysis was conducted using RLK sequences from plants other than *Arabidopsis*. The sequences analyzed all fell into the same clade as *Arabidopsis* RLKs (data not shown).

Among the ePK families found in *Arabidopsis*, Raf kinases were paraphyletic to the RLK family and, together with RLKs, formed a well supported group with a bootstrap value of 98% (Fig. 2). Based on the parsimony criterion, the support for the RLK family and Raf kinases as a monophyletic group was still high at 86% (data not shown). Taken together, these results indicated that Raf kinases are the closest relatives to RLKs among the *Arabidopsis* sequences analyzed.

The Relationships Between Animal Receptor Kinases and Plant RLKs. Animal RTKs and receptor serine/threonine kinases (RSKs) are other families of ePKs with a domain organization similar to that of the plant RLKs. To determine the relationships among these receptor kinase families, we analyzed the phylogenetic relationships between the kinase domain sequences of representative *Arabidopsis* RLKs and animal receptor kinases. *Arabidopsis* and human representatives of other ePK families were also included.

EVOLUTION

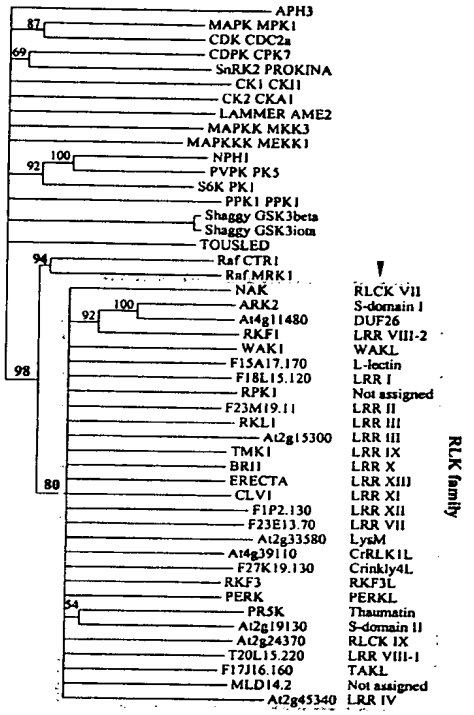


Fig. 2. *Arabidopsis* receptor-like kinases and related kinases form a monophyletic group distinct from all of the other eukaryotic protein kinases found in the *Arabidopsis* genome. The tree was generated with the kinase domain sequences of representative *Arabidopsis* ePKs and RLKs using APH(3')III as outgroup based on minimal evolution. The bootstrap values are shown at the nodes. The boxed region represents the receptor-like kinase family. The arrowhead indicates the RLK subfamily. The abbreviations used are according to Fig. 1.

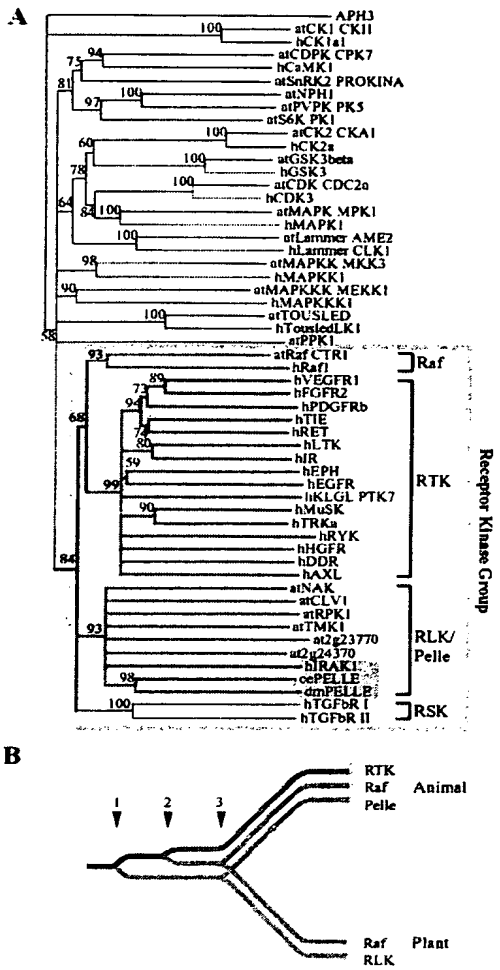


Fig. 3. Human receptor kinases and *Arabidopsis* receptor kinases belong to distinct but related families, and Pelle kinases are the animal homologs of *Arabidopsis* RLKs. (*A*) Plant and animal representatives of ePKs were used in this analysis. The tree is rooted with APH(3')III based on minimal evolution. It indicates that Raf, RSK, RTK, and RLK form a well supported group distinct from all other ePKs (boxed region). The bootstrap values are shown at the nodes. Animal Pelle kinases (shaded area) are found in the same clade as RLKs. (*B*) The proposed evolutionary relationships between receptor kinase family members are as follows: 1, an ancient duplication event leading to the divergence of RLK/Pelle from RTK/Raf; 2, a more recent gene duplication leading to the divergence of RTK from Raf; and 3, the divergence of plant and animal lineages, resulting in the ancestral sequences that gave rise to the extant receptors and related kinases.

The phylogenetic tree generated based on minimal evolution criterion is shown in Fig. 3. All 16 RTK subfamily representatives and c-SRC formed a well supported group, indicating a monophyletic origin for tyrosine receptor kinases. The sister groups to the RTK family were Raf kinases. Plant RLKs included in this analysis formed another monophyletic group, indicating that RLKs have a distinct origin from that of Raf kinases and animal RTKs. Plant RLKs, Raf kinases, RSKs, and RTKs collectively formed a well supported group with a bootstrap support of 84%. The monophyly of kinases in this group when compared with the other ePK families was also supported by analyses based on maximum parsimony (data not shown). However, the specific relationships between animal RSKs, RTKs, Raf kinases, and plant RLKs were less conclusive because different optimality criteria gave inconsistent results. To investigate whether the results obtained were biased by using only human sequences, we conducted an extended analysis including RTK and RSK sequences from *Caenorhabditis*, *Drosophila*, sponge, and hydra, and we reached the same conclusion (data not shown). Based on these analyses, we defined the monophyletic group that contains the RLK, RTK, RSK, and Raf kinase genes as the receptor kinase group.

Homologs of Plant RLKs in Eukaryotes. To determine whether members of the RLK family are present in organisms other than flowering plants, we first used the kinase domain sequence of CLV1 to search for homologous sequences in the genomes of yeast, *C. elegans*, *D. melanogaster*, and human. No RLK homolog was found in the yeast genome. Five animal homologs of the RLK family were found: the Pelle kinase (DmPelle) in *Drosophila* (40), the Pelle-like kinase (CePelle) in *Caenorhabditis* (T23534), and three IRAKs in human (32). DmPelle, CePelle, and IRAK1 are all cytoplasmic kinases and all fell into the same clade as plant RLKs with strong bootstrap support (Fig. 3, shaded area). A similar search using other RLK kinase sequences yielded the same results (data not shown). Based on this

analysis, we defined the clade containing the plant RLKs and Pelle-like sequences as the RLK/Pelle family.

To broaden the scope of the searches, we used the amino acid sequences of CLV1 and Pelle kinase domains to search the EST database for RLK homologs in 20 different eukaryotes (15 shown in Table 1). Sequences with an E value of less than 1×10^{-50}, a conservative criterion, were regarded as RLK homologs without further examination. The remaining sequences were subjected to BLAST searches and were treated as RLK homologs if the top five matches were known RLKs or RLK homologs. The results of EST searches are shown in Table 1. All seven of the flowering plants, including four dicots and three monocots, have 0.18% to 0.6% of their ESTs representing RLK/Pelle family members. Pines, ferns, and mosses all have a percent EST representation similar to that of flowering plants. With the exception of the three ESTs representing *Drosophila* Pelle kinase, no other organism examined produced ESTs, which could be classified as RLK/Pelle family members.

The Distribution of RLKs on *Arabidopsis* Chromosomes. The size discrepancy of the RLK/Pelle family between plants and animals raises the question on how the expansion of this family occurred in the plant lineages. To address this question, we examined the location of RLKs on the *Arabidopsis* chromosomes. After comparing the location of genes to the phylogeny based on kinase domains, we found that subfamilies differed in their chromosomal distributions. At one extreme, 35 of the 40 members of the DUF26 subfamily were located on chromosome 4 (Fig. 4*A*). At the other extreme, 51 genes representing LRR X, XI, and XIII subfamilies were distributed among all five chromosomes (Fig. 4*B*). In addition, we found that more than 30% of the RLK/Pelle family members in *Arabidopsis* are in tandem repeats with 2 to 19 genes. A closer look at the location of the 38 DUF26 subfamily members on chromosome 4 (including three potential pseudogenes) indicates that 34 of them are in tandem repeats (Fig. 4*C*). The phylogenetic relationships between DUF26 genes in the tandem repeats indicates that at least one intrachromosomal duplication event occurred in the region containing tandem repeats. Taken together, the results suggest that tandem duplication events and large-scale duplications of chromosomes are two of the potential mechanisms responsible for the expansion of the RLK/Pelle family in *Arabidopsis*.

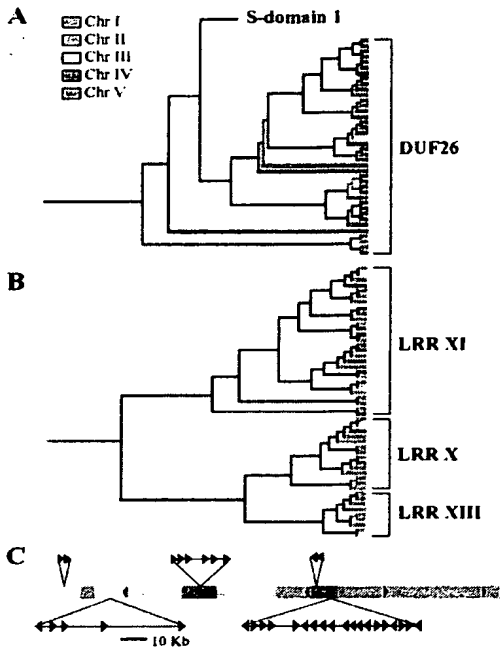


Fig. 4. Distribution of RLKs on *Arabidopsis* chromosomes provides clues for the mechanisms of RLK family expansion. (*A*) The cladogram of the DUF26 subfamily was generated with the kinase domain sequences based on minimal evolution criterion. The color coding on branches indicates the chromosome on which each gene in the subfamily is located. Note that most DUF26 members are located on chromosome 4. (*B*) The cladogram of LRR X, XI, and XIII subfamilies was generated and color-coded in the same manner as *A*. Note that most genes derived from duplication events are located on different chromosomes. (*C*) A detailed depiction of DUF26 distribution on chromosome 4 indicates that tandem duplications and an internal chromosomal duplication may contribute extensively to the expansion of this subfamily. The 10-kb legend is for the expanded region showing tandem repeats. The regions with postulated chromosomal duplications are color-coded according to their similarity to regions on the other chromosomes. The color-coding scheme is the same as *A*. Three potential DUF26 pseudogenes are also included in the diagram.

Discussion

Evolutionary History of the Receptor Kinase Group. Plant RLKs were originally grouped into a single family based on their configuration as transmembrane kinases with serine and threonine specificity. Our analysis provides a phylogenetic basis for the classification of RLKs as a single family in the eukaryotic protein kinase superfamily. Interestingly, 24% of the 610 *Arabidopsis* genes in the RLK/Pelle family analyzed do not have an extracellular domain based on the absence of signal sequences and transmembrane regions. Some of these apparently cytoplasmic kinases form unique subfamilies, whereas others are most closely related to kinases with a receptor topology. The latter may represent ancestral forms that were recruited into the receptor kinase configuration by domain fusion events. Alternatively, some of the soluble kinase forms could be derived from ancestral receptor kinase forms. In any case, it is apparent that kinase domains from the RLK/Pelle family were recruited multiple times by fusion with different extracellular domains to form the subfamilies found in *Arabidopsis*. This notion can be expanded to include the animal RTK and RSK families in the receptor kinase group, which appear to have been formed by recruitment of kinases from the same lineage, distinct from all other ePK families.

Based on the kinase domain phylogeny, a hypothetical sequence of events that occurred in the evolution of the receptor kinase group is proposed in Fig. 3*B*. According to this model, an early gene duplication event led to the founding of two lineages that diversified into the RTK and Raf families on one hand and the RLK/Pelle family on the other. This diversification seems to have occurred before the divergence of plants and animals. In addition, both lineages contain representatives of soluble kinase and transmembrane receptor forms. It should be noted that the soluble Pelle-like and Raf kinases form complexes with cell surface receptors and are responsible for transduction of signals to downstream effectors (33, 34). Perhaps the continual recruitment of this particular lineage of kinase modules was favored during evolution because ancestral forms had already specialized in mediating signaling from transmembrane receptors. Examination of kinases belonging to the receptor kinase group in more primitive eukaryotes may be informative. Whereas fungi such as yeast and *Neurospora* do not appear to have representatives of the receptor kinase group, the slime mold, *Dictyostelium discoideum*, has several examples (data not shown). None of these sequences from slime mold has predicted signal peptide or transmembrane regions, and most of the sequences are dual specificity kinases based on their kinase activities (35), consistent with the possibility that the ancestral form for extant receptor kinases may have been soluble kinases.

Diversification of the Plant RLK/Pelle Family. The small number of representatives of the RLK/Pelle family in animals compared with the much larger number in *Arabidopsis* indicates that the expansion of the plant RLKs occurred after the divergence of plant and animal lineages or that massive gene loss occurred in the animals. A comparison of EST representation with the known total number of RLK/Pelle members in the fully sequenced genomes of *C. elegans*, *D. melanogaster*, and *Arabidopsis* indicated that the EST representation provided a conservative estimate of the total number of family members in the genomes. The lack of RLK/Pelle ESTs in *Porphyra* and *Chlamydomonas* argues that, rather than massive gene loss in the animal genomes examined, the RLK/Pelle family likely underwent expansion after the divergence of animal and plant lineages. Interestingly, all land plants have similar percent representations of RLK/Pelle kinases, suggesting that the size of this gene family may have been similar to the present-day level before the diversification of the land plant lineages. Additional sequence information will be necessary to determine whether all RLK subfamilies found in *Arabidopsis* are equally represented in these other land plant lineages. The early expansion of the RLK/Pelle family could be associated with evolution of multicellularity, as has been suggested for the RTK family in animals (36). Alternatively, the expansion of the family could be associated with the development of the complex array of attributes required for the migration of plant lineages from the aquatic to the terrestrial environment. Examination of RLK/Pelle representation in multicellular green algae such as *Chara* could help to resolve this question.

The monophyletic origin of the RLK/Pelle family implies that the expansion of the family to its present size in *Arabidopsis* was the result of multiple gene-duplication events. Two possible mechanisms for the amplification of this family are suggested by the way members of some subfamilies are distributed on the *Arabidopsis* chromosomes. For example, the DUF26 subfamily is organized in tandem arrays (Fig. 4*C*). These tandem arrays were likely generated by gene duplications resulting from unequal crossing-over as seen in the other gene families such as disease resistance genes (37). Gene duplication is also driven by larger scale duplication events, including polyploidization followed by reshuffling of chromosomal regions (5, 38, 39). Tandem arrays of DUF26 members are located in such duplicated regions on chromosome 4. However, the localization of other DUF26 subfamily members almost exclusively on chromosome 4 suggests that this subfamily expanded after the extensive chromosome duplications and reshuffling identified for multiple regions of all five *Arabidopsis* chromosomes (5, 38, 39).

On the other hand, members of LRR X, XI, and XIII subfamilies are distributed among all five chromosomes, with related genes on each branch of the phylogenetic tree generally located on different chromosomes. These three related subfamilies are of particular interest because they include *Arabidopsis* RLKs with known developmental functions such as BRI1, CLV1, ERECTA, and HAESA (6, 7, 10). The difference in distribution patterns between the DUF26 and these LRR subfamilies could indicate that the LRR subfamilies originally expanded by mechanisms that did not include localized (e.g., tandem) duplications. Given sufficient evolutionary time, several rounds of polyploidization followed by chromosomal rearrangements could produce a given subfamily of the observed size from a single prototypical gene. Alternatively, these LRR subfamilies may have originally expanded via localized duplications that occurred early enough in evolutionary time that extensive chromosome reshuffling could have eliminated linkage between subfamily members. Both proposed mechanisms imply that the LRR X, XI, and XIII subfamilies may have expanded much earlier in time than the DUF26 subfamily. A comparative analysis of RLK subfamilies in other plant lineages should help to resolve this issue.

We thank Micahel Gribskov for advice on sequence retrieval and alignments and Donna Fernandez, Frans Tax, Sara E. Patterson, and Melissa D. Lehti-Shiu for reading the manuscript. This work was supported by Department of Energy Grant DE-FG02-91ER20029 (to A.B.B.) and National Research Initiative Competitive Grants Program/U.S. Department of Agriculture Grant 2000-21469 (to S.-H.S.).

1. van der Geer, P., Hunter, T. & Lundberg, R. A. (1994) *Annu. Rev. Cell Biol.* **10**, 251–337.
2. Robertson, S. C., Tynan, J. A. & Donoghue, D. J. (2000) *Trends Genet.* **16**, 265–271.
3. Walker, J. C. (1994) *Plant Mol. Biol.* **26**, 1599–1609.
4. McCarty, D. R. & Chory, J. (2000) *Cell* **103**, 201–209.
5. Arabidopsis Genome Initiative. (2000) *Nature (London)* **408**, 796–815.
6. Li, J. & Chory, J. (1997) *Cell* **90**, 929–938.
7. Clark, S. E., Williams, R. W. & Meyerowitz, E. M. (1997) *Cell* **89**, 575–585.
8. Gomez-Gomez, L. & Boller, T. (2000) *Mol. Cell* **5**, 1003–1011.
9. Becraft, P. W., Stinard, P. S. & McCarty, D. R. (1996) *Science* **273**, 1406–1409.
10. Jinn, T. L., Stone, J. M. & Walker, J. C. (2000) *Genes Dev.* **14**, 108–117.
11. Stein, J. C., Howlett, B., Boyes, D. C., Nasrallah, M. E. & Nasrallah, J. B. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8816–8820.
12. Song, W. Y., Wang, G. L., Chen, L. L., Kim, H. S., Pi, L. Y., Holsten, T., Gadner, J., Wang, B., Zhai, W. X., Zhu, L. H., *et al.* (1995) *Science* **270**, 1804–1806.
13. Schopfer, C., Nasrallah, M. & Nasrallah, J. (1999) *Science* **286**, 1697–1700.
14. Takayama, S., Shiba, H., Iwano, M., Shimosato, H., Che, F.-S., Kai, N., Watanabe, M., Suzuki, G., Hinata, K. & Isogai, A. (1999) *Proc. Natl. Acad. Sci. USA* **97**, 1920–1925.
15. Brand, U., Fletcher, J. C., Hobe, M., Meyerowitz, E. M. & Simon, R. (2000) *Science* **289**, 617–619.
16. Trotochaud, A. E., Jeong, S. & Clark, S. E. (2000) *Science* **289**, 613–617.
17. Wang, Z.-Y., Seto, H., Fujioka, S., Yoshida, S. & Chory, J. (2001) *Nature (London)* **410**, 380–383.
18. Gomez-Gomez, L., Bauer, Z. & Boller, T. (2001) *Plant Cell* **13**, 1155–1163.
19. Stone, J. M., Collinge, M. A., Smith, R. D., Horn, M. A. & Walker, J. C. (1994) *Science* **266**, 793–795.
20. Bower, M. S., Matias, D. D., Fernandes-Carvalho, E., Mazzurco, M., Gu, T., Rothstein, S. J. & Goring, D. R. (1996) *Plant Cell* **8**, 1647–1650.
21. Trotochaud, A. E., Hao, T., Wu, G., Yang, Z. & Clark, S. E. (1999) *Plant Cell* **11**, 393–406.
22. Hanks, S. K. & Hunter, T. (1995) *FASEB J.* **9**, 576–596.
23. Hardie, D. G. (1999) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 97–131.
24. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389–3402.
25. Schultz, J., Copley, R. R., Doerks, T., Ponting, C. P. & Bork, P. (2000) *Nucleic Acids Res.* **28**, 231–234.
26. Sonnhammer, E. L. L., Eddy, S. R., Birney, E., Bateman, A. & Durbin, R. (1998) *Nucleic Acids Res.* **26**, 320–322.
27. Higgins, D. G., Thompson, J. D. & Gibson, T. J. (1996) *Methods Enzymol.* **266**, 383–402.
28. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406–425.
29. Swafford, D. L. (1998) PAUP*: Phylogenetic Analysis Using Parsimony (*and other materials) (Sinauer, Sunderland, MA).
30. Hon, W.-C., McKay, G. A., Thompson, P. R., Sweet, R. M., Yang, D. S. C., Wright, G. D. & Berghuis, A. M. (1997) *Cell* **89**, 887–895.
31. Leonard, C. J., Aravind, L. & Koonin, E. V. (1998) *Genome Res.* **8**, 1038–1047.
32. Cao, Z., Henzel, W. J. & Gao, X. (1996) *Science* **271**, 1128–1131.
33. O'Neil, L. A. & Greene, C. (1998) *J. Leukocyte Biol.* **63**, 650–657.
34. Daum, G., Eisenmann-Tappe, I., Fries, H.-W., Troppmair, J. & Rapp, U. R. (1994) *Trends Biochem. Sci.* **19**, 474–480.
35. Nuckolls, G. H., Osherov, N., Loomis, W. F. & Spudich, J. A. (1996) *Development (Cambridge, U.K.)* **122**, 3295–3305.
36. Hunter, T. & Plowman, G. D. (1997) *Trends Biochem. Sci.* **22**, 18–22.
37. Ronald, P. C. (1998) *Curr. Opin. Plant Biol.* **1**, 294–298.
38. Blanc, G., Barakat, A., Guyot, R., Cooke, R. & Delseny, M. (2000) *Plant Cell* **12**, 1093–1101.
39. Vision, T. J., Brown, D. G. & Tanksley, S. D. (2000) *Science* **290**, 2114–2117.
40. Shelton, C. A. & Wasserman, S. A. (1993) *Cell* **72**, 515–525.

EXHIBIT C

Academy Translations
a business of Tenco Services Pty Ltd
ABN 72 892 315 097

PO Box 259
Kyneton VIC 3444
Australia

Free: 1800 637 640
Tel: (03) 5423 2558
Fax: (03) 8625 0041

www.academyXL.com
info@academyXL.com

TRANSLATION VERIFICATION CERTIFICATE

This is to certify that the attached document is an **English translation** of the

– Japanese Patent Application 9-56382 –

and *Academy Translations* declare that the translation thereof is to the best of their knowledge and ability true and correct.

December 23, 2008

. . . . . .

Date

Stamp/Signature:

Academy Translations
PO Box 259, Kyneton VIC 3444 AUSTRALIA

AT Ref.: sf-2227

Applicant: Josette Masle et al.
Serial No.: 10/519,135
Filed: August 15, 2005
Exhibit C

Gene Coding for Protein Controlling Morphogenesis of Plants

0001

Relevant area of industry

This invention relates to genetic DNA that controls the morphogenesis of plants, and to DNA that codes for antisense RNA for such genes, and to plants whose morphology is transformed by such DNA, and provides a technique for adjusting the stem lengths and inflorescent forms of plants.

0002

Prior art

It is believed that plant morphology is affected by genetic factors and environmental factors. Hitherto, one method of creating short stemmed plants and plants of different inflorescent forms was to vary the morphology of useful varieties by genetic hybridization with plants of different morphology, but it was difficult to consistently obtain individuals with superior characteristics to those of the new variety.

0003

Moreover, variations occur in genes affecting morphology through sudden spontaneous variation and through induced sudden variation and individuals can be selected that exhibit morphological change through reduced genetic function, but it was very difficult to deliberately

create individuals in which variations occurred which retained unaltered the useful acquired characteristics simply through genes that controlled specific morphogenesis.

0004

Thus, it is believed that, were it possible to isolate the genes that control plant morphogenesis, it would be possible to create plants with reduced stem lengths by transforming the plant morphology by incorporating in the genes a vector that would express antisense RNA (antisense RNA expression vector).

0005

In *Shiroi Nunazuna* (*Arabidopsis thaliana*), elongation of the floral buds accompanies the change from the vegetative stage to the reproductive stage, and the growth of the buds is strongly associated with the elongation of the subsequent internodes, displaying a highly ordered branching pattern. The *Landsberg erecta* strain which is known as the standard ecotype of *Arabidopsis thaliana* retains the endogenous erecta variation and exhibits different floral buds. The flowers form dense, compact floral buds in the crown. The variation is pleiotropic and possesses round leaves and short flat siliques (the above is an abstract published at the UCLA Keystone Symposium, Huang, I. et al., (1991)).

0006

Strains possessing variations at the same gene locus as the variation of the *Landsberg erecta* strain have yielded

dwarf variations known to genetics, which have been named the er-101 strain, er-102 strain and er-103 strain.

0007

Problem to be solved by the invention

It has been shown that a correlation between the dwarfing of plants and specific genes is to a certain extent known, but the gene itself has not yet been found.

0008

The present invention was developed from this knowledge, and it is an objective of the invention to provide a gene that controls plant morphogenesis. Moreover it is an objective of the invention to provide plants that possess stems of greater or less length, or plants in which the inflorescence is altered.

0009

Means employed in order to solve the problem

In order to achieve this objective, the inventors of the present invention cloned the gene that controls the morphogenesis of the plant from the chromosomal DNA of *Arabidopsis thaliana*, and having obtained an antisense expression vector in which the gene was combined, discovered that the morphology of the individual plants whose morphology had been transformed by the vector DNA was altered and thus arrived at the present invention.

0010

Thus the present invention is of DNA that codes for a protein that includes an amino acid sequence that possesses the amino acid sequence shown in Sequence Number 1 that possesses the action of controlling plant

morphogenesis or alternatively the sequence in which one or a plurality of amino acid residues that do not affect the action of controlling plant morphogenesis are substituted, excised or inserted.

0011

Moreover, the invention provides DNA that codes for antisense RNA that controls the expression of the aforementioned DNA. This DNA is further characterized by possessing a substantively complementary base sequence in at least one portion of the base sequence in Sequence Number 1.

0012

The inventors furthermore provide plants whose morphology has been transformed by the DNA that codes for the aforementioned protein, and also provide plants whose morphology has been transformed by DNA that codes for the aforementioned antisense RNA.

0013

The protein coded for by the DNA of the present invention is a protein related to the control of the morphogenesis of plants, having extensive expression on the stems and flowers thereof, which are greatly altered, and in particular to the elongation of the stems, the gene that codes for the protein transforming the morphology of plants, and being expected to promote the elongation of stems through the increased expression level of the gene.

0014

On the other hand, it may also be anticipated that the elongation of the stem of the morphologically transformed

plant could be controlled through the morphological transformation of the plant by the DNA that codes for the aforementioned antisense RNA, or in other words, the DNA that controls the expression of the gene related to the control of the morphogenesis of the plant.

0015

In these Specifications, 'chromosomal DNA' and 'chromosomal gene' refer to the DNA included in the nuclear chromosomes of the plant cells and the gene present on such DNA. Moreover, the gene that relates to the control of the morphogenesis of plants provided by the present invention may be referred to as the 'morphogene' or the 'gene envisaged by the present invention'.

0016

Mode of implementation of the invention

The mode of implementation of the invention is described below. The gene envisaged by the present invention may be acquired from plants that possess the variation relating to the variation of expressed morphology by the isolation of the gene relating to such variation of expressed morphology. Moreover, wild-type genes may be acquired from the chromosomal DNA of wild-type plants through hybridization with the oligonucleotides prepared on the basis of the base sequences of the acquired variant gene as a probe, or with pairs of oligonucleotides prepared on the basis of the base sequences of the variant gene as the primer by polymerase chain reaction (PCR).

0017

The preparation of plants that possess the variation relating to the control of morphogenesis, the method of isolating the gene of the present invention from the variant, and the method of use of the gene are now described in detail. The general methods required for gene recombination, such as DNA incision, junction and transformation, and the determination of gene base sequences and hybridization and so forth, are described in the documentation accompanying the commercial enzymes used in these operations and in Molecular Cloning (Maniatis, T. et al., Cold Spring Harbor Laboratory Press).

0018

<1> *Isolation and identification of gene controlling plant morphogenesis*

(1) Preparation of plants possessing variation relating to the control of morphogenesis

The mutagenesis method (genetic disruption) method in which the gene is disrupted at the insertion site through the introduction of foreign genes into the plant cell and the insertion of chromosomal DNA is employed in order to cause variation in the gene that controls morphogenesis in plants such as *Arabidopsis thaliana*. Methods of gene introduction that may be employed are the use of *Agrobacterium*, electropolation of plant protoplastic cells, the polyethylene glycol method and microinjection and so forth. The use of *Agrobacterium* is the most effective of these methods for reasons of high transformation efficiency in *Arabidopsis thaliana* and because few variations apart from the variations due to the gene introduction are caused.

0019

Here foreign genes are introduced by the agrobacterium infection method employing a binary vector system (vector system including a functional replication point with T-DNA or *E. coli* or other microorganisms which are capable of introducing DNA into plant cells and preferably a marker gene for selecting the plant or microorganism cells) into the plant cells and plants are prepared with variations in the gene that controls morphogenesis.

0020

Wild strains of *Arabidopsis thaliana* are infected by the *in planta* agrobacterium infection method (Chang, S.S., Park, S.K. et al.: *Plant J.*, 5, No.4 (1994)) and so forth with Ti plasmid-derived binary vectors and are harvested after growth for approximately 6 weeks. The resulting seeds are implanted in agar-agar cultures containing hygromycin and are cultivated. Those transformed plants that exhibit hygromycin resistance are transplanted into rock wool, and the plants are selected by visual inspection for variants (*erecta* variants) whose stem lengths are different from those of those of the wild plants.

0021

The variants acquired in this manner possess the strong possibility of variation through the insertion of the binary vector into the gene that controls plant morphogenesis.

0022

(2) Isolation of variant gene from morphogenes

The variant genes relating to this variation are isolated by what is known as the plasmid rescue method from the variants exhibiting variant morphology acquired in this manner. Thus chromosomal DNA is prepared by the method described in *Cell*, 35 (1983), p.35 from the variants and is excised by means of restriction enzymes and the ends within the molecules are linked by self ligation. If the resulting ring DNA includes the binary vector into which chromosomes have been inserted, the DNA molecules function as plasmids capable of autonomous replication in *E. coli* cells and the transformed individuals exhibit resistance to markers (such as Ampicillin).

0023

The ring DNA is transformed by means of *E. coli* and the recombinant plasmids are recovered from the transformed individuals that are resistant to markers, and chromosomal DNA segments that include the binary vector and morphogene can be acquired.

0024

Alternatively, chromosomal DNA is prepared from the variant plants and is excised by means of a suitable restriction enzyme and is then linked to a plasmid or phage vector, and a chromosome library is prepared through transformation of *E. coli*. Clones are selected from the library with T-DNA and the like as probes, and clones possessing variant gene fragments can be selected.

0025

(3) Isolation of morphogenes from wild type genes

Chromosomal DNA libraries are prepared from wild type *Arabidopsis thaliana* by employing P1 phage vectors and

the like, and the wild type morphogenes can be isolated by hybridization with the variant gene fragments that have been acquired in this manner as the probes. Primers are prepared on the basis of the base sequences of the variant genes, and the gene envisaged by the present invention can also be acquired by amplification of the wild type genes from the chromosomal DNA of the wild type plants by PCR.

0026

The base sequence of the DNA fragments containing the morphogene of *Arabidopsis thaliana* derived from the practical embodiments described below is shown in Sequence Table Sequence Number 2. This gene contains 27 exons and 26 introns.

0027

The gene envisaged by the present invention as explained above contains many introns. The cDNA of the morphogenes can be isolated in order to acquire the exon part, that is the DNA that codes for the protein that controls morphogenesis. The cDNA library can be created by extracting the mRNA from the terrestrial tissue of *Arabidopsis thaliana*, synthesis of the DNA by means of a reverse transcription enzyme, insertion of a two-chained part into a vector by means of a polymerase reaction, and transformation of *E. coli* and the like. cDNA cloning kits are commercially available and may be employed. The morphogene cDNA clones are obtained with the chromosomal genes as probes from the resulting cDNA libraries.

0028

The base sequence of the cDNA from the practical embodiment described below, and the amino acid sequence inferred from the base sequence, are shown as Sequence Table Sequence Number 1. The translated product of this gene exhibits similarity to RLK5 shown in *Nature*, 345 (1990), p.743. Gene RLK5 was isolated as a receptor-like protein kinase present in cell membranes, but because it was isolated through its similarity to the base sequences of known protein kinase genes, the gene and its translated product are completely unknown. The pattern of expression of the RLK5 gene is different from that of the gene controlling morphogenesis and the gene is expressed in the roots as well as the terrestrial part, and it is believed that the two genes are functionally different.

0029

<2> Use of morphogene

The gene envisaged by the present invention is a gene associated with the control of plant morphogenesis, and particularly with the elongation of the stem, and it is anticipated that the elongation of plant stems can be promoted through the transformation of plants with the gene and an increase in the level of expression of the gene.

0030

In order to use the gene envisaged by the present invention to transform plants, DNA may be introduced into the protoplast by electropolation or alternatively by using the Ti plasmid of *Agrobacterium*. In this case, chromosomal genes or cDNA prepared from mRNA may be employed as the genes envisaged by the present invention.

0031

On the other hand, it is anticipated that it will be possible to control the elongation of the stems of transformed plants by transforming plants with DNA that expresses antisense RNA that controls the expression of the gene envisaged by the present invention, that is, RNA that possesses a complementary sequence over the entire length or at least in part of the mRNA transcribed from the morphogene.

0032

The DNA that expresses the antisense RNA is acquired by linking the antisense chain (chain possessing a complementary base sequence to the sense chain (code chain)) or at least portion thereof to the downstream of the promoter. In other words, two DNA chains that include sequences that are similar to the code chain or at least to part thereof are acquired by linking to the downstream of the promoter in the reverse direction to the direction of the original transcription. The antisense chain is acquired from the chromosomal DNA or cDNA, but the use of the exon portion is preferred because the introns cannot be anticipated to exercise the function of controlling the expression of the gene envisaged by the present invention when chromosomal DNA is used. Moreover, any of the code region, the 5′ untranslated region or the 3′ untranslated region may be used as at least portion of the antisense chain. Moreover, the DNA that codes for the antisense RNA may include a sequence that codes for complementary poly-dU in the poly(A) chain added to the 3′ end of the mRNA in addition to the 3′ untranslated region.

0033

CaMV 35S promoter may be used as the promoter used in the present invention. This may be used in the same manner as the transformation of plants by the gene envisaged by the present invention in the transformation of plants by transforming plants by DNA that codes for antisense RNA.

0034

Furthermore, the upstream of the coding region for the gene envisaged by the present invention includes a region that controls the expression of the gene envisaged by the present invention. This region is a region including at least the sequences represented by Base Numbers 1 to 1752 in Sequence Number 2, and more specifically, the region of Base Numbers 396 to 1752. This region is anticipated to be capable of use in the control of the expression of genes in plant cells.

0035

Practical embodiment

The present invention is described in greater detail below by means of a practical embodiment of the invention.

(1) Preparation of variant plants in relation to the control of morphogenesis

Shiroi Nunazuna (*Arabidopsis thaliana*) ecotype WS (Wassilewskija, purchased from Lehle Seeds) was infected *in planta* with *Agrobacterium* strain EHA101 (*Agrobacterium tumefaciens*) that possessed pGDW32 binary vector (possessing hygromycin resistance gene and ampicillin resistance gene, capable of autonomous replication in *E.*

coli cells) derived from Ti plasmids by the agrobacterium infection method (Chang, S.S., Park, S.K. et al.: *Plant J.*, 5, No.4 (1994)).

0036

This is described more specifically below. *Agrobacterium* into which pGDW32 had been introduced was proliferated overnight in an LB culture containing the antibiotic (hygromycin). 10 μl of this culture was diluted (1/20) with 190 μl of Gamborg's B5 culture medium (containing 2% sucrose, pH 5.5). After from 19 to 23 days from inoculation, agrobacterium treatment was performed at the growth stage at which bolting had only commenced.

0037

Only those flower stems that had begun to lengthen were cut at their bases with a #11 scalpel, and a 26Gx1/2 size hypodermic syringe was employed to perforate the rosette stem from the incision. 1 μl of dilute *Agrobacterium* solution was injected into the wound. During this time, the plants were irradiated with 3000 to 4000 lux. Three days after inoculation, the plants were transplanted into rock wool fibre mini-pots and were then grown in the normal manner. Approximately 6 weeks after inoculation, the seeds were harvested from the rapidly formed approximately half siliques.

0038

The seeds obtained were implanted in B5 agar-agar culture containing 10 μg/ml of hygromycin and were grown under a 12 hours daylight/12 hours night lighting cycle under 6,000 to 12,000 lux of irradiation at a temperature of

22° C and at a humidity of 30% to 40%. HYPONeX (manufactured by Murakami Bussan KK) diluted to 1/1000 was used as the nutrient. The transformed plants that possessed pGDW32 T-DNA inserted sequences that exhibited hygromycin resistance were transplanted to rock wool and grown, and the T4 seeds (fourth generation seeds) were collected. The plants were inspected visually at this time, and the variants that exhibited different elongations of the stems from the wild type plants were acquired.

0039

(2) Isolation of variant genes

Arabidopsis genome DNA was prepared by the method described in *Cell*, 35 (1983), p.35.

0040

5 g of *Arabidopsis* tissue (root) was ground to a fine powder in a mortar in liquid nitrogen at -80° C this was added to 25 ml of DNA isolation buffer (50 mM tris-HCl, pH 7.5, 0.2 M NaCl, 20 mM EDTA-Na2, 2% N-lauroyl sarcosine sodium salt, 3 g/ml urea and 5% TE saturated phenol) and stirred, 25 ml of phenol/chloroform was added, whereupon 1.5 ml of 10% SDS (sodium dodecyl sulphate) was added and the mixture was stirred gently at room temperature for 10 minutes. This was centrifuged for 10 minutes at 6000 rpm, whereupon 25 ml of phenol/chloroform was added to the aqueous layer and the solution was centrifuged again for 10 minutes at 6000 rpm. 15 ml of phenol was added to the aqueous layer and the solution was stirred and was then immediately centrifuged for 10 minutes at 6000 rpm. The supernatant was discarded, 25 ml of 70% ethanol was added to the

sediment, the mixture was stirred in a Vortex and was then centrifuged for 10 minutes at 6000 rpm, whereupon the supernatant was discarded and the residue was dried under reduced pressure. The residue was dissolved in 400 μl of TE at a pH of 8.0 (10 μg/μl RNase).

0041

200 μg of chromosomal DNA prepared in this manner from variants with control of morphogenesis was then refined by the CsCl ultracentrifugation method. 1 μg of this chromosomal DNA was excised by EcoRI or XbaI, the restriction enzyme fragments were refined by phenol extraction and ethanol precipitation and the ends within the molecules were linked by self-ligation; *E. coli* (XL1-Blue MRF' (purchased from Stratagene)) was then transformed with the resulting ring DNA. This yielded approximately one thousand colonies of a transformed strain that exhibited resistance to ampicillin. Plasmid DNA was extracted from the resistant colonies acquired and analysis was performed.

0042

Those of the rescued plasmids that had been excised from the DNA by EcoRI were labelled pREa and pREb, and those that had been excised from the DNA by XbaI were labelled pRXa and pRXb.

0043

(3) Isolation of variant genes

The variant genes were isolated from the chromosomal DNA library with the chromosomal DNA fragments containing the plasmids rescued in the manner described above as the

probes. P1 phage vector (purchased from Du Pont) was employed in order to prepare a nuclear DNA library of the *Arabidopsis* ecotype Colombia strain by the method described in *The Plant Journal*, 7 (1995), p.351. The EcoRI and XbaI fragments prepared respectively from pRXb and pREa were ^{32}P labeled and were used as probes in order to perform plaque hybridization. Two positive clones were obtained as a result; these were labeled 28D7 and 61H10 respectively. Restriction enzyme maps were prepared for these two clones, and it was found that 28D7 held approximately 25 kb and 61H10 held approximately 75 kb of inserted fragments derived from *Arabidopsis* chromosomal DNA (Figure 1).

0044

Subcloning of the variant genes contained in the inserted DNA was performed in order to determine the base sequences. Restriction enzyme maps were prepared of the sequences adjacent to the T-DNA and Southern analyses were performed of the inserted fragments in order to estimate the T-DNA insertion sites on the inserted fragments (Figure 2). The chromosomal DNA sequences in the vicinity of the insertion sites were subcloned in Bluescript II SK+. The sites on the inserted fragments of the resulting subclones are shown in Figure 3.

0045

(4) Isolation of cDNA clones of genes related to the control of morphogenesis

mRNA was prepared from the terrestrial portion tissue of the *Arabidopsis* ecotype Colombia strain, and cDNA was prepared by the molecular cloning method (Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989): A Laboratory

Manual, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), and a cDNA library was prepared by using λYES (purchased from Clontech) as the vector. The sequences adjacent to the T-DNA were excised from the aforementioned pRXb and pREa and were ^{32}P labeled. These were used as the probes when plaque hybridization was performed on 300,000 plaques of the phage library. The selected plasmid was pKUT161.

0046

(5) Analysis of gene related to the control of morphogenesis

The base sequences of the cDNA clone contained in pKUT161 and the chromosomal gene were analyzed. The amino acid sequence was inferred from the cDNA base sequence and this sequence is shown in Sequence Table Sequence Number 1. The base sequence of the DNA fragment including the chromosomal gene is shown in Sequence Number 2.

0047

When the analysis of the amino acid sequence inferred from the cDNA base sequence was performed, the translated product of the gene was found to exhibit similarities to the RLK5 described in *Nature*, 345 (1990) p.743. The RLK5 gene was isolated as a receptor-like protein kinase present on the cell membrane, but because it was isolated through its similarity to the base sequences of known protein kinase genes, the gene and the translated products thereof are entirely unknown. The pattern of expression of the RLK5 gene is different from the gene related to the control of morphogenesis, the gene is expressed not only in the terrestrial part but also in

the roots, and the genes are considered by be functionally different genes.

0048

(6) Analysis of variant related to control of morphogenesis

The variant related to the control of morphogenesis acquired at (1) was considered from the elongation of the stem to be a variant of the same gene locus as the erecta variant of the *Landsberg erecta* strain. Thus first the genetic complementarity of the *Landsberg erecta* strain, the er-103 strain and the variant isolated by the present invention was tested. As a result, it was found that all these variants arose from a variation in the same gene locus and the variant isolated was the er-104 strain.

0049

mRNA was isolated from the *Landsberg erecta* strain and er-103 strain and was amplified by the RT-PCR (reverse transcription - PCR) method. Thus the mRNA formed the template and the reverse transcription reaction was performed with the oligonucleotide possessing the base sequence shown in Sequence Number 4 as the primer; the cDNA acquired was then used as the template, and the PCR reaction was performed with the aforementioned oligonucleotide as the 3′ end primer and an oligonucleotide possessing the base sequence shown in Sequence Number 3 as the 5′ end primer.

0050

The resulting amplified product was used as a template and the base sequence was determined by the direct sequence method using the aforementioned oligonucleotide

primer. As a result it was found that in the *Landsberg erecta* strain, the isoleucine residue had been substituted by a lysine residue in association with the substitution of A for T in Base Number 2299 in Sequence Number 1 of amino acid No. 750. In the er-103 strain, the methionine residue had been substituted by an isoleucine residue in Amino Acid Number 282 due to the substitution of A for G in Base Number 896 in Sequence Table Sequence Number 1.

0051

It was found that in all these variants, portions of the base sequence for the gene related to the control of morphogenesis possessed variations that gave rise to changes in the amino acid sequences, and this gene could be identified as the gene that controlled morphogenesis (morphogene).

0052

(7) Analysis of expression of morphogene

As a further method of confirming that the cDNA isolated in the manner described above and the chromosomal gene was the gene that controlled morphogenesis, analyses were made of the extent of expression in each of the variant plant tissues, and Northern Analysis was employed in order to make a comparative analysis of the extent of gene expression between the wild type plants and the variant type plants. cDNA clones were used as the probes. As a result of the Northern Analysis using whole individual wild type plants and *Landsberg erecta* variants, the expression of the morphogene (formation of mRNA) was observed as strong in the wild type plants, but was barely detected in the variants.

0053

Moreover, as a result of analyses of the extent of expression of the morphogene in each of the tissues of the wild type plant, the extent of expression was high in variants in which variation in the stems and flowers and so forth was great, and this also confirmed that the gene was the gene related to the elongation of the stem.

0054

Moreover, as the expression of the morphogene is site and time specific as shown above, the expression control region for the gene envisaged by the present invention can be used to control the expression of exogenous genes in plants. The expression control region is a region that includes at least portion of the sequence represented by Base Numbers 1 to 1752 in Sequence Number 2, and more specifically, the region of Base Numbers 396 to 1752.

0055

Effects of the invention

The invention provides a gene that controls the morphogenesis of plants. It can be anticipated that the elongation of the stems of plants could be promoted by increasing the extent of expression of the gene. Moreover, it could be anticipated that the elongation of the stems of plants could be controlled by transforming plants by means of the DNA sequence that expresses the antisense RNA for the gene.

0056

Sequence Tables

Sequence Number: 1

Sequence length: 3176

Sequence type: Nucleic acid

Number of chains: Two

Topology: Straight chain

Class of sequence: cDNA t mRNA

Source

Plant name: Shiroi Nunazuna (*Arabidopsis thaliana*)

Strain: Colombia

Characteristics of sequence:

Symbol indicating characteristics: CDS

Sites: 51..2978

```
CTTTTAAAGT ATATCTAAAA ACGCAGTCGT TTTAAGACTG TGTGTGAGAA ATG GCT      56
                                                          Met Ala
                                                            1

CTG TTT AGA GAT ATT GTT CTT CTT GGG TTT CTC TTC TGC TTG AGC TTA     104
Leu Phe Arg Asp Ile Val Leu Leu Gly Phe Leu Phe Cys Leu Ser Leu
  5                 10                  15
```

```
GTA GCT ACT GTG ACT TCA GAG GAG GGA GCA ACG TTG CTG GAG ATT AAG      152
Val Ala Thr Val Thr Ser Glu Glu Gly Ala Thr Leu Leu Glu Ile Lys
    20                  25                  30
AAG TCA TTC AAA GAT GTG AAC AAT GTT CTT TAT GAC TGG ACA ACT TCA      200
Lys Ser Phe Lys Asp Val Asn Asn Val Leu Tyr Asp Trp Thr Thr Ser
35                  40                  45                  50
CCT TCT TCG GAT TAT TGT GTC TGG AGA GGT GTG TCT TGT GAA AAT GTC      248
Pro Ser Ser Asp Tyr Cys Val Trp Arg Gly Val Ser Cys Glu Asn Val
                55                  60                  65
ACC TTC AAT GTT GTT GCT CTT AAT TTG TCA GAT TTG AAT CTT GAT GGA      296
Thr Phe Asn Val Val Ala Leu Asn Leu Ser Asp Leu Asn Leu Asp Gly
            70                  75                  80
GAA ATC TCA CCT GCT ATT GGA GAT CTC AAG AGT CTC TTG TCA ATT GAT      344
Glu Ile Ser Pro Ala Ile Gly Asp Leu Lys Ser Leu Leu Ser Ile Asp
        85                  90                  95
CTG CGA GGT AAT CGC TTG TCT GGA CAA ATC CCT GAT GAG ATT GGT GAC      392
Leu Arg Gly Asn Arg Leu Ser Gly Gln Ile Pro Asp Glu Ile Gly Asp
    100                 105                 110
TGT TCT TCT TTG CAA AAC TTA GAC TTA TCC TTC AAT GAA TTA AGT GGT      440
Cys Ser Ser Leu Gln Asn Leu Asp Leu Ser Phe Asn Glu Leu Ser Gly
115                 120                 125                 130
GAC ATA CCG TTT TCG ATT TCG AAG TTG AAG CAA CTT GAG CAG CTG ATT      488
Asp Ile Pro Phe Ser Ile Ser Lys Leu Lys Gln Leu Glu Gln Leu Ile
                135                 140                 145
CTG AAG AAT AAC CAA TTG ATA GGA CCG ATC CCT TCA ACA CTT TCA CAG      536
Leu Lys Asn Asn Gln Leu Ile Gly Pro Ile Pro Ser Thr Leu Ser Gln
            150                 155                 160
ATT CCA AAC CTG AAA ATT CTG GAC TTG GCA CAG AAT AAA CTC AGT GGT      584
Ile Pro Asn Leu Lys Ile Leu Asp Leu Ala Gln Asn Lys Leu Ser Gly
        165                 170                 175
GAG ATA CCA AGA CTT ATT TAC TGG AAT GAA GTT CTT CAG TAT CTT GGG      632
Glu Ile Pro Arg Leu Ile Tyr Trp Asn Glu Val Leu Gln Tyr Leu Gly
    180                 185                 190
TTG CGA GGA AAC AAC TTA GTC GGT AAC ATT TCT CCA GAT TTG TGT CAA      680
Leu Arg Gly Asn Asn Leu Val Gly Asn Ile Ser Pro Asp Leu Cys Gln
195                 200                 205                 210
CTG ACT GGT CTT TGG TAT TTT GAC GTA AGA AAC AAC AGT TTG ACT GGT      728
Leu Thr Gly Leu Trp Tyr Phe Asp Val Arg Asn Asn Ser Leu Thr Gly
                215                 220                 225
AGT ATA CCT GAG ACG ATA GGA AAT TGC ACT GCC TTC CAG GTT TTG GAC      776
Ser Ile Pro Glu Thr Ile Gly Asn Cys Thr Ala Phe Gln Val Leu Asp
            230                 235                 240
TTG TCC TAC AAT CAG CTA ACT GGT GAG ATC CCT TTT GAC ATC GGC TTC      824
Leu Ser Tyr Asn Gln Leu Thr Gly Glu Ile Pro Phe Asp Ile Gly Phe
        245                 250                 255
CTG CAA GTT GCA ACA TTA TCA TTG CAA GGC AAT CAA CTC TCT GGG AAG      872
Leu Gln Val Ala Thr Leu Ser Leu Gln Gly Asn Gln Leu Ser Gly Lys
    260                 265                 270
ATT CCA TCA GTG ATT GGT CTC ATG CAA GCC CTT GCA GTC TTA GAT CTA      920
Ile Pro Ser Val Ile Gly Leu Met Gln Ala Leu Ala Val Leu Asp Leu
```

```
275                 280                 285                 290
AGT GGC AAC TTG TTG AGT GGA TCT ATT CCT CCG ATT CTC GGA AAT CTT      968
Ser Gly Asn Leu Leu Ser Gly Ser Ile Pro Pro Ile Leu Gly Asn Leu
                295                 300                 305
ACT TTC ACC GAG AAA TTG TAT TTG CAC AGT AAC AAG CTG ACT GGT TCA     1016
Thr Phe Thr Glu Lys Leu Tyr Leu His Ser Asn Lys Leu Thr Gly Ser
            310                 315                 320
ATT CCA CCT GAG CTT GGA AAC ATG TCA AAA CTC CAT TAC CTG GAA CTC     1064
Ile Pro Pro Glu Leu Gly Asn Met Ser Lys Leu His Tyr Leu Glu Leu
        325                 330                 335
AAT GAT AAT CAT CTC ACG GGT CAT ATA CCA CCA GAG CTT GGG AAG CTT     1112
Asn Asp Asn His Leu Thr Gly His Ile Pro Pro Glu Leu Gly Lys Leu
    340                 345                 350
ACT GAC TTG TTT GAT CTG AAT GTG GCC AAC AAT GAT CTG GAA GGA CCT     1160
Thr Asp Leu Phe Asp Leu Asn Val Ala Asn Asn Asp Leu Glu Gly Pro
355                 360                 365                 370
ATA CCT GAT CAT CTG AGC TCT TGC ACA AAT CTA AAC AGC TTA AAT GTT     1208
Ile Pro Asp His Leu Ser Ser Cys Thr Asn Leu Asn Ser Leu Asn Val
                375                 380                 385
CAT GGG AAC AAG TTT AGT GGC ACT ATA CCC CGA GCA TTT CAA AAG CTA     1256
His Gly Asn Lys Phe Ser Gly Thr Ile Pro Arg Ala Phe Gln Lys Leu
            390                 395                 400
GAA AGT ATG ACT TAC CTT AAT CTG TCC AGC AAC AAT ATC AAA GGT CCA     1304
Glu Ser Met Thr Tyr Leu Asn Leu Ser Ser Asn Asn Ile Lys Gly Pro
        405                 410                 415
ATC CCG GTT GAG CTA TCT CGT ATC GGT AAC TTA GAT ACA TTG GAT CTT     1352
Ile Pro Val Glu Leu Ser Arg Ile Gly Asn Leu Asp Thr Leu Asp Leu
    420                 425                 430
TCC AAC AAC AAG ATA AAT GGA ATC ATT CCT TCT TCC CTT GGT GAT TTG     1400
Ser Asn Asn Lys Ile Asn Gly Ile Ile Pro Ser Ser Leu Gly Asp Leu
435                 440                 445                 450
GAG CAT CTT CTC AAG ATG AAC TTG AGT AGA AAT CAT ATA ACT GGT GTA     1448
Glu His Leu Leu Lys Met Asn Leu Ser Arg Asn His Ile Thr Gly Val
                455                 460                 465
GTT CCA GGC GAC TTT GGA AAT CTA AGA AGC ATC ATG GAA ATA GAT CTT     1496
Val Pro Gly Asp Phe Gly Asn Leu Arg Ser Ile Met Glu Ile Asp Leu
            470                 475                 480
TCA AAT AAT GAT ATC TCT GGC CCA ATT CCA GAA GAG CTT AAC CAA TTA     1544
Ser Asn Asn Asp Ile Ser Gly Pro Ile Pro Glu Glu Leu Asn Gln Leu
        485                 490                 495
CAG AAC ATA ATT TTG CTG AGA CTG GAA AAT AAT AAC CTG ACT GGT AAT     1592
Gln Asn Ile Ile Leu Leu Arg Leu Glu Asn Asn Asn Leu Thr Gly Asn
    500                 505                 510
GTT GGT TCA TTA GCC AAC TGT CTC AGT CTC ACT GTA TTG AAT GTA TCT     1640
Val Gly Ser Leu Ala Asn Cys Leu Ser Leu Thr Val Leu Asn Val Ser
515                 520                 525                 530
CAT AAC AAC CTC GTA GGT GAT ATC CCT AAG AAC AAT AAC TTC TCA AGA     1688
His Asn Asn Leu Val Gly Asp Ile Pro Lys Asn Asn Asn Phe Ser Arg
                535                 540                 545
TTT TCA CCA GAC AGC TTC ATT GGC AAT CCT GGT CTT TGC GGT AGT TGG     1736
```

```
Phe Ser Pro Asp Ser Phe Ile Gly Asn Pro Gly Leu Cys Gly Ser Trp
          550                 555                 560
CTA AAC TCA CCG TGT CAT GAT TCT CGT CGA ACT GTA CGA GTG TCA ATC    1784
Leu Asn Ser Pro Cys His Asp Ser Arg Arg Thr Val Arg Val Ser Ile
       565                 570                 575
TCT AGA GCA GCT ATT CTT GGA ATA GCT ATT GGG GGA CTT GTG ATC CTT    1832
Ser Arg Ala Ala Ile Leu Gly Ile Ala Ile Gly Gly Leu Val Ile Leu
    580                 585                 590
CTC ATG GTC TTA ATA GCA GCT TGC CGA CCG CAT AAT CCT CCT CCT TTT    1880
Leu Met Val Leu Ile Ala Ala Cys Arg Pro His Asn Pro Pro Pro Phe
595                 600                 605                 610
CTT GAT GGA TCA CTT GAC AAA CCA GTA ACT TAT TCG ACA CCG AAG CTC    1928
Leu Asp Gly Ser Leu Asp Lys Pro Val Thr Tyr Ser Thr Pro Lys Leu
                615                 620                 625
GTC ATC CTT CAT ATG AAC ATG GCA CTC CAC GTT TAC GAG GAT ATC ATG    1976
Val Ile Leu His Met Asn Met Ala Leu His Val Tyr Glu Asp Ile Met
          630                 635                 640
AGA ATG ACA GAG AAT CTA AGT GAG AAG TAT ATC ATT GGG CAC GGA GCA    2024
Arg Met Thr Glu Asn Leu Ser Glu Lys Tyr Ile Ile Gly His Gly Ala
       645                 650                 655
TCA AGC ACT GTA TAC AAA TGT GTT TTG AAG AAT TGT AAA CCG GTT GCG    2072
Ser Ser Thr Val Tyr Lys Cys Val Leu Lys Asn Cys Lys Pro Val Ala
    660                 665                 670
ATT AAG CGG CTT TAC TCT CAC AAC CCA CAG TCA ATG AAA CAG TTT GAA    2120
Ile Lys Arg Leu Tyr Ser His Asn Pro Gln Ser Met Lys Gln Phe Glu
675                 680                 685                 690
ACA GAA CTC GAG ATG CTA AGT AGC ATC AAG CAC AGA AAT CTT GTG AGC    2168
Thr Glu Leu Glu Met Leu Ser Ser Ile Lys His Arg Asn Leu Val Ser
                695                 700                 705
CTA CAA GCT TAT TCC CTC TCT CAC TTG GGG AGT CTT CTG TTC TAT GAC    2216
Leu Gln Ala Tyr Ser Leu Ser His Leu Gly Ser Leu Leu Phe Tyr Asp
          710                 715                 720
TAT TTG GAA AAT GGT AGC CTC TGG GAT CTT CTT CAT GGC CCT ACG AAG    2264
Tyr Leu Glu Asn Gly Ser Leu Trp Asp Leu Leu His Gly Pro Thr Lys
       725                 730                 735
AAA AAG ACT CTT GAT TGG GAC ACA CGG CTT AAG ATA GCA TAT GGT GCA    2312
Lys Lys Thr Leu Asp Trp Asp Thr Arg Leu Lys Ile Ala Tyr Gly Ala
    740                 745                 750
GCA CAA GGT TTA GCT TAT CTA CAC CAT GAC TGT AGT CCA AGG ATC ATT    2360
Ala Gln Gly Leu Ala Tyr Leu His His Asp Cys Ser Pro Arg Ile Ile
755                 760                 765                 770
CAC AGA GAC GTG AAG TCG TCC AAC ATT CTC TTG GAC AAA GAC TTA GAG    2408
His Arg Asp Val Lys Ser Ser Asn Ile Leu Leu Asp Lys Asp Leu Glu
                775                 780                 785
GCT CGT TTG ACA GAT TTT GGA ATA GCG AAA AGC TTG TGT GTG TCA AAG    2456
Ala Arg Leu Thr Asp Phe Gly Ile Ala Lys Ser Leu Cys Val Ser Lys
          790                 795                 800
TCA CAT ACT TCA ACT TAC GTG ATG GGC ACG ATA GGT TAC ATA GAC CCC    2504
Ser His Thr Ser Thr Tyr Val Met Gly Thr Ile Gly Tyr Ile Asp Pro
       805                 810                 815
```

```
GAG TAT GCT CGC ACT TCA CGG CTC ACT GAG AAA TCC GAT GTC TAC AGT      2552
Glu Tyr Ala Arg Thr Ser Arg Leu Thr Glu Lys Ser Asp Val Tyr Ser
    820                 825                 830
TAT GGA ATA GTC CTT CTT GAG CTG TTA ACC CGA AGG AAA GCC GTT GAT      2600
Tyr Gly Ile Val Leu Leu Glu Leu Leu Thr Arg Arg Lys Ala Val Asp
835                 840                 845                 850
GAC GAA TCC AAT CTC CAC CAT CTG ATA ATG TCA AAG ACG GGG AAC AAT      2648
Asp Glu Ser Asn Leu His His Leu Ile Met Ser Lys Thr Gly Asn Asn
                855                 860                 865
GAA GTG ATG GAA ATG GCA GAT CCA GAC ATC ACA TCG ACG TGT AAA GAT      2696
Glu Val Met Glu Met Ala Asp Pro Asp Ile Thr Ser Thr Cys Lys Asp
            870                 875                 880
CTC GGT GTG GTG AAG AAA GTT TTC CAA CTG GCA CTC CTA TGC ACC AAA      2744
Leu Gly Val Val Lys Lys Val Phe Gln Leu Ala Leu Leu Cys Thr Lys
        885                 890                 895
AGA CAG CCG AAT GAT CGA CCC ACA ATG CAC CAG GTG ACT CGT GTT CTC      2792
Arg Gln Pro Asn Asp Arg Pro Thr Met His Gln Val Thr Arg Val Leu
    900                 905                 910
GGC AGT TTT ATG CTA TCG GAA CAA CCA CCT GCT GCG ACT GAC ACG TCA      2840
Gly Ser Phe Met Leu Ser Glu Gln Pro Pro Ala Ala Thr Asp Thr Ser
915                 920                 925                 930
GCG ACG CTG GCT GGT TCG TGC TAC GTC GAT GAG TAT GCA AAT CTC AAG      2888
Ala Thr Leu Ala Gly Ser Cys Tyr Val Asp Glu Tyr Ala Asn Leu Lys
                935                 940                 945
ACT CCT CAT TCT GTC AAT TGC TCT TCC ATG AGT GCT TCT GAT GCT CAA      2936
Thr Pro His Ser Val Asn Cys Ser Ser Met Ser Ala Ser Asp Ala Gln
            950                 955                 960
CTG TTT CTT CGG TTT GGA CAA GTT ATT TCT CAG AAC AGT GAG              2978
Leu Phe Leu Arg Phe Gly Gln Val Ile Ser Gln Asn Ser Glu
        965                 970                 975
TAGTTTTTCG TTAGGAGGAG AATCTTTAAA ACGGTATCTT TTCGTTGCGT TAAGCTGTTA   3038
GAAAAATTAA TGTCTCATGT AAAGTATTAT GCACTGCCTT ATTATTATTA GACAAGTGTG   3098
TGGTGTGAAT ATGTCTTCAG ACTGGCACTT AGACTTCCAA AAAAAAAAAA AAAAAAAAAA   3158
AAAAAAAAAA AAAAAAAA                                                 3176
```

0057

Sequence Number 2

Sequence length: 9295

Sequence type: Nucleic acid

Number of chains: Two

Topology: Straight chain

Class of sequence: Genomic DNA

Source

Plant name: Shiroi Nunazuna (*Arabidopsis thaliana*)

Strain: Colombia
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 1803..1881
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 1882..2227
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 2228..2366
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 2367..2467
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 2540..2643
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 2468..2539
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 2644..2715
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 2716..2809
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 2810..2878
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 2879..2968
Characteristics of sequence:
Symbol indicating characteristics: Exon

Sites: 2969..3040
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 3041..3118
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 3119..3190
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 3191..3266
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 3267..3338
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 3339..3421
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 3422..3493
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 3494..3586
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 3587..3655
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 3656..3740
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 3741..3812
Characteristics of sequence:
Symbol indicating characteristics: Intron

Sites: 3813..3888
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 3889..3960
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 3961..4048
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 4049..4120
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 4121..4209
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 4210..4281
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 4282..4349
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 4350..4421
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 4422..4508
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 4509..4580
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 4581..4706
Characteristics of sequence:
Symbol indicating characteristics: Exon

Sites: 4707..4778
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 4779..4860
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 4861..4932
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 4933..5018
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 5019..5090
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 5091..5176
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 5177..5248
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 5249..5412
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 5413..5481
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 5482..5576
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 5577..5648
Characteristics of sequence:
Symbol indicating characteristics: Intron

Sites: 5649..5726
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 5727..5800
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 5801..5882
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 5883..6011
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 6096..6443
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 6012..6095
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 6444..6519
Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 6520..6890
Characteristics of sequence:
Symbol indicating characteristics: Intron
Sites: 6891..6974 Characteristics of sequence:
Symbol indicating characteristics: Exon
Sites: 6975..7328

Sequence

GAATTCAAAG GAATAAGCAT CGGAGACGAT TTAATGTTAC CTCTTGACGT ATTTATCCAA 60
TTTATCCATT AAGCCACCAG CCATAGCATC TGATCATCAT CATCAACATA TAAATAACCA 120
AATTTGAAAT GAACAAAAGT CGAATTGGTG ATATTGAAAA TCGAGTTCGT GAAATTGAGA 180
ATCGGATTGG TGAATTTGAA GAGAGATGCG TGTACCGTTA GGGAGGAGGA GGAGACGGGA 240
GAGAAAAAAG GAGACGGAGA TAACTCGCCG GCTCTGTTTC CATGGCGGAG GTGATAATGT 300
AGCTGCGCAC GTTAGCTTTT TGTGGTTTGA GTTGGAGAAC AGTGGGAGGC TCACGGTAGC 360
GTGGAGTGAC GACATTGGGG ATAACACCAG AGGCGTCTTA TCTCCGTTGG ACAAATTATT 420
ATTATGGCTA TGAACATTCA ACATATAATT TAATTAGAAA TTTGCGGATG AAAAAGAGGT 480
AAACAATTGC AGAAATGGTT AAAAATATTA ACGTTGTACA GCAAATGATA ATAAAAAGTG 540
TAACGTACAG TGTGTAAGGA ATGGAAAAAT AATAATTTGG GTTAAAATAA ATATGTAGTT 600
TTCTAACTAT ATAGTACTTT TTGAGAAAAG ATAATATTAT GTGTATTTTT ATTGAAACAA 660
ATAAATGATT TAACAAAAAA AAAAAGAGAA GTTAAAATGA AAAGGAATTA TTATTTTTTA 720
AGTTCTTCCT TCTTTTGTTG GGCCTGTGAC CCTTTTAGTT TTAGTCCACT TCGTTCTCAA 780
AGCTTCAAAA TATTAATTTT GTGACAAACC GACCGGAGCC AACCAAACCG GTTAACATCC 840
TAAAACCAAT CATATTTTAT TAAGTTTTGT GTTGATGCTA AACCAAAAAT CATTGGCATG 900
CATATTTCTA AATTTAGTAA TAAACAAAAA CACTTAGAAA TCACACGTTC ACTATACTAA 960
AAAACGTTGA CAAAAACACA ACAACTATAC TAATAATTAA AGAAGAGAAA ACTGAACCAA 1020
ACTTTTTGTA AACTCCTGAA TTTAAATTAG TAATTGAAGT AAGAAGATGA AGAAGAACAT 1080
GTTAAGCAAA CAAAAAAATT ACACTAAAAT CATATAAAAA TACATAATTA CAAAAGTACC 1140
CATAAGATGG ATTTATTGAT ATGGGTCATC TGTGAAACAA GCCACAGAGA GACAAAGACT 1200
CGTAAGTATT GGGCAACGAA AGCGACCTCC TTTATTCACC ACTGCCATTA ACATGTTCTT 1260
CTTCTCCTTC TTCTTCTACA TTTTATGACC GTTTTACCCT TCAAGAGAGA GAAACAAAAT 1320
CACTCCCTCT CACTCACTCT ATCTCTCTCT TCTGCAAAGC TTCAGAACTC TGGCAGAGAG 1380
ATAAAAGATG ATGGGGTTTT TAACTTTATC CTCCCCAAAT AATTCTTCTT CCCTTCATCT 1440
CTCTCTCTTA CACAACAGGT CCCTACATTT GTACAATCTC CTCTCTTTAA AGACTCTCTC 1500
TCTTTCTCTC TCCATCTCTA TCTTACTCTG TATTTCTGTC GTCTGAGCAC TCAATGAAAC 1560
CACTGTAAAT TTCCGCCAGA ATTTGATGTG ATGGAACGAT AAAAATCATT TTTTCTCGGT 1620
TAAAGTAAAA AAACAAAAAC AAATTTCTGT AGAAATCATA ATAAAAGAAA GAAAAAAAAT 1680
CTAATGTCGG TACATAATAC GGTTCTCTTC TTCTTCTCTA TCCTCTGTTT CTTCTTCATG 1740
GAGACTTGAA AGCTTTTAAA GTATATCTAA AAACGCAGTC GTTTTAAGAC TGTGTGTGAG 1800
AAATGGCTCT GTTTAGAGAT ATTGTTCTTC TTGGGTTTCT CTTCTGCTTG AGCTTAGTAG 1860

CTACTGTGAC TTCAGAGGAG GGTCAGTTAT TATACTGATG CATGCTTCTT CAAGTTCAAG 1920
ATTTTCGTCT TTTTGTTTTA TATTAGTGAA AAAAACTTAA AGATGAGATT TTTATATGAT 1980
TTTTGAAGTT TCATTTGGTG AAAATGAGAT CTGGGTACTT GTTATTTTCT ATTTTTGCTT 2040
TTTGTAATGG TTTTTTTTTA CTTGGTGGGT CTTCTATAGA ATCAAAAGAA GCTTTGAATA 2100
AATTAGGGTT TGAGTTTTAT TTTGTTTTCT TGGAAGTTGA ATTTTTAATC TTCTCAAGAA 2160
CTGACAAATA TTTTTTTTTG TTTTTGTGCG TGTGTGTTAA TAAAATATCC TTAAAACAAA 2220
ATTAAAGGAG CAACGTTGCT GGAGATTAAG AAGTCATTCA AAGATGTGAA CAATGTTCTT 2280
TATGACTGGA CAACTTCACC TTCTTCGGAT TATTGTGTCT GGAGAGGTGT GTCTTGTGAA 2340
AATGTCACCT TCAATGTTGT TGCTCTGTAA GTTTCTTCAT TCCTTTAGAT TACTATTACA 2400
GTGGTTTTTG GTGTTCTTGT GGGAAAAAGT TGTAATTTGT TTTGTGTGTG TTTTCTATGT 2460
TTTGTAGTAA TTTGTCAGAT TTGAATCTTG ATGGAGAAAT CTCACCTGCT ATTGGAGATC 2520
TCAAGAGTCT CTTGTCAATG TAACTGTTTC AACATTCACT GTAGCATGAA ATAAAGTATC 2580
TTACTTTAAT TCTATTCCAC TCTCTGAGTT GTGACTTTTG TCTTCTGTTT TTTTCTAATG 2640
TAGTGATCTG CGAGGTAATC GCTTGTCTGG ACAAATCCCT GATGAGATTG GTGACTGTTC 2700
TTCTTTGCAA AACTTGTAAG AACAGTGATT GGTGTTATTC TACCATTAAA CTTTTGTTCA 2760
TAGAGGTTTT ATTTGATGAA GTGTGTTCAT GTTGTTTTTA ATTCAGAGAC TTATCCTTCA 2820
ATGAATTAAG TGGTGACATA CCGTTTTCGA TTTCGAAGTT GAAGCAACTT GAGCAGCTGT 2880
AAGTAGCTAG TTATTCTGCT ACTAGTCTTC ATATGTCATT GCTAAAAATA TACTCACCAT 2940
GTGGAATATG GATTTTTACT TTGTCCAGGA TTCTGAAGAA TAACCAATTG ATAGGACCGA 3000
TCCCTTCAAC ACTTTCACAG ATTCCAAACC TGAAAATTCT GTATGTTCCC CATGATTCTT 3060
ACATGTCTTA CTACTTTTAG CTATATAGGT GATCATACAT GTGTAATTTC AATTGCAGGG 3120
ACTTGGCACA GAATAAACTC AGTGGTGAGA TACCAAGACT TATTTACTGG AATGAAGTTC 3180
TTCAGTATCT GTAAGTGTCA ATGTTTTTTG AAGTCTGTCA ATGTCTCTTC ATTACCCGGT 3240
GATAATTGTT GTACTATGAT GAGCAGTGGG TTGCGAGGAA ACAACTTAGT CGGTAACATT 3300
TCTCCAGATT TGTGTCAACT GACTGGTCTT TGGTATTTGT GAGTCTTCTT GCACATCTGA 3360
ATAGTATGAT GAGTTCTTTT GTAAATATCA AATATCTGAC TTTGTTTTGA TATTGAATCA 3420
GTGACGTAAG AAACAACAGT TTGACTGGTA GTATACCTGA GACGATAGGA AATTGCACTG 3480
CCTTCCAGGT TTTGTATGTG CCTCTTTCTC TACTTCTAAA CATCATTACT GTAATTTGGG 3540
TTACTTAAGA AAATCTACTT AACTGGTTTG CTTATTACGA ACTCAGGGAC TTGTCCTACA 3600
ATCAGCTAAC TGGTGAGATC CCTTTTGACA TCGGCTTCCT GCAAGTTGCA ACATTGTTAG 3660
TTCTCACCTC TACTAATCTT TTGCTTTAAA TTTTGGCTAG CCTTTGTTTT CTTTTAAAGA 3720
AGATCATTTT CTTATCTTAG ATCATTGCAA GGCAATCAAC TCTCTGGGAA GATTCCATCA 3780
GTGATTGGTC TCATGCAAGC CCTTGCAGTC TTGTAAGTAC TTTTCTTCTA ATCAATGAAG 3840
CTACTTATAA CATTTTCATG AACTTAGGTT ATATGTTTTC TTTTACAGAG ATCTAAGTGG 3900
CAACTTGTTG AGTGGATCTA TTCCTCCGAT TCTCGGAAAT CTTACTTTCA CCGAGAAATT 3960
GTAATTCTTT ACCTGTTTGT TTTCAGTTTG GAGTCAAATG TCATACCATG TTAATGATAG 4020
TGATTTATCT TTTTGGCTTT ATCTCTAGGT ATTTGCACAG TAACAAGCTG ACTGGTTCAA 4080
TTCCACCTGA GCTTGGAAAC ATGTCAAAAC TCCATTACCT GTATGACCAA CCTTCTCTTC 4140
ACTTCTCTTT TTGCATACAG TCACTACTAA GTTGTGTTTC CTTATCAACT ATTTGTAAAA 4200
TATTCATAGG GAACTCAATG ATAATCATCT CACGGGTCAT ATACCACCAG AGCTTGGGAA 4260
GCTTACTGAC TTGTTTGATC TGTAAGTAGT TCTTCCTATG CTTGACATGT TTTGATGTTC 4320
TTATGCTTAT ATGAACTATG TACATATAGG AATGTGGCCA ACAATGATCT GGAAGGACCT 4380
ATACCTGATC ATCTGAGCTC TTGCACAAAT CTAAACAGCT TGTATGTATC TCTTTCTCTG 4440
AAAACTTCTC ACTTGAATGT TCAAGATTGG TGCTTTATAT GATTTTGTGT CTCATTAATG 4500
TAATGTAGAA ATGTTCATGG GAACAAGTTT AGTGGCACTA TACCCCGAGC ATTTCAAAAG 4560
CTAGAAAGTA TGACTTACCT GTAAGTATCG ACGCTGAGAA TTTCTCTAAT CTTATATAAT 4620
ATATAGTTCC ACAGCGTTTG TTTTTTCGAA TTTCAAGTCA TTAACTACTG AGTTTTTGGT 4680
TGCCTTTGAT TTATCGGTTC AACCAGTAAT CTGTCCAGCA ACAATATCAA AGGTCCAATC 4740
CCGGTTGAGC TATCTCGTAT CGGTAACTTA GATACATTGT AAGTGTTTCT TGTTTTCTGT 4800
GAAGTATACA TCATTATATG TGCCTTGTCT CACATTTATT AAATTTAATG ACATTTGAAG 4860

GGATCTTTCC AACAACAAGA TAAATGGAAT CATTCCTTCT TCCCTTGGTG ATTTGGAGCA 4920
TCTTCTCAAG ATGTGAGCAT CCATAAGACC TCCAGTTTTA TTGTTTATTT CTAGCAAAAG 4980
ATGAAAATGG TTTGTGAACT CTTGCATTCT TGTTATAGGA ACTTGAGTAG AAATCATATA 5040
ACTGGTGTAG TTCCAGGCGA CTTTGGAAAT CTAAGAAGCA TCATGGAAAT GTAAGAAGTT 5100
AACTTCTATC TGCTTGGTTA GAGTTTTTTT CATTTATCTC AATTACTGTT CTGAATTTGT 5160
GTGTTTGTGG TTGCAGAGAT CTTTCAAATA ATGATATCTC TGGCCCAATT CCAGAAGAGC 5220
TTAACCAATT ACAGAACATA ATTTTGCTGT AAGCAATCTT CCTCTTATCC CTTCCAAGCT 5280
GTTAAGAAAT TGTTTTTGTA GAATGAAACT AAAACTCTGT ATACACAATA ATGAGGTCAC 5340
TATAGTGTGA TCCAGGAACA TGTATTGGGT TGGTGATCTA TCTAATGTTG TGTTTCTTAA 5400
AATTGCTTGC AGGAGACTGG AAAATAATAA CCTGACTGGT AATGTTGGTT CATTAGCCAA 5460
CTGTCTCAGT CTCACTGTAT TGTAAGTAGG CACCTTTGGT TCTGAAACAT TTTTTGTCCC 5520
TCTTTGTGCA TCTTTTGCTA AGAATATAAC CCTGCAATCT TCACTAACTC TTATAGGAAT 5580
GTATCTCATA ACAACCTCGT AGGTGATATC CCTAAGAACA ATAACTTCTC AAGATTTTCA 5640
CCAGACAGGT ATGGTAATTT AGCAGGTTTT GGTATTGTGC ATTTTGTTTT GTTTGCTAAT 5700
ATCTATGTTT ATGTTTTTGG ATAAAGCTTC ATTGGCAATC CTGGTCTTTG CGGTAGTTGG 5760
CTAAACTCAC CGTGTCATGA TTCTCGTCGA ACTGTACGAG GTGATTACAT TCTTCTAAAA 5820
GCTTCCATTC ACAAAACCTA AGATAATTAA AGCTCATGTT TCTATCCATG TTTTGTCTGC 5880
AGTGTCAATC TCTAGAGCAG CTATTCTTGG AATAGCTATT GGGGGACTTG TGATCCTTCT 5940
CATGGTCTTA ATAGCAGCTT GCCGACCGCA TAATCCTCCT CCTTTTCTTG ATGGATCACT 6000
TGACAAACCA GGTCTACTCT CCAAACCACT TTACGAATGT TCTTCACCTA CAATGTAATC 6060
CAATAGTTAA TCCTTAAATT TCCTGGTGAC ATCAGTAACT TATTCGACAC CGAAGCTCGT 6120
CATCCTTCAT ATGAACATGG CACTCCACGT TTACGAGGAT ATCATGAGAA TGACAGAGAA 6180
TCTAAGTGAG AAGTATATCA TTGGGCACGG AGCATCAAGC ACTGTATACA AATGTGTTTT 6240
GAAGAATTGT AAACCGGTTG CGATTAAGCG GCTTTACTCT CACAACCCAC AGTCAATGAA 6300
ACAGTTTGAA ACAGAACTCG AGATGCTAAG TAGCATCAAG CACAGAAATC TTGTGAGCCT 6360
ACAAGCTTAT TCCCTCTCTC ACTTGGGGAG TCTTCTGTTC TATGACTATT TGGAAAATGG 6420
TAGCCTCTGG GATCTTCTTC ATGGTAAGTC TCATCGCCAA ACATAGAAAA TTATTTGAAT 6480
CTTCTGTGAC ATAACAACTT GCTTGTGTGT TTTGTAAAGG CCCTACGAAG AAAAAGACTC 6540
TTGATTGGGA CACACGGCTT AAGATAGCAT ATGGTGCAGC ACAAGGTTTA GCTTATCTAC 6600
ACCATGACTG TAGTCCAAGG ATCATTCACA GAGACGTGAA GTCGTCCAAC ATTCTCTTGG 6660
ACAAAGACTT AGAGGCTCGT TTGACAGATT TTGGAATAGC GAAAAGCTTG TGTGTGTCAA 6720
AGTCACATAC TTCAACTTAC GTGATGGGCA CGATAGGTTA CATAGACCCC GAGTATGCTC 6780
GCACTTCACG GCTCACTGAG AAATCCGATG TCTACAGTTA TGGAATAGTC CTTCTTGAGT 6840
TGTTAACCCG AAGGAAAGCC GTTGATGACG AATCCAATCT CCACCATCTG GTTTGTTCTT 6900
TCTTGCCTAT CTCTCTCAGC TGCTCTGTTT AGGTCAAGTC CGTAATCTTG TTTTCATTGA 6960
TTCACTTACA TCAGATAATG TCAAAGACGG GGAACAATGA AGTGATGGAA ATGGCAGATC 7020
CAGACATCAC ATCGACGTGT AAAGATCTCG GTGTGGTGAA GAAAGTTTTC CAACTGGCAC 7080
TCCTATGCAC CAAAAGACAG CCGAATGATC GACCCACAAT GCACCAGGTG ACTCGTGTTC 7140
TCGGCAGTTT TATGCTATCG GAACAACCAC CTGCTGCGAC TGACACGTCA GCGACGCTGG 7200
CTGGTTCGTG CTACGTCGAT GAGTATGCAA ATCTCAAGAC TCCTCATTCT GTCAATTGCT 7260
CTTCCATGAG TGCTTCTGAT GCTCAACTGT TTCTTCGGTT TGGACAAGTT ATTTCTCAGA 7320
ACAGTGAGTA GTTTTTCGTT AGGAGGAGAA TCTTTAAAAC GGTATCTTTT CGTTGCGTTA 7380
AGCTGTTAGA AAAATTAATG TCTCATGTAA AGTATTATGC ACTGCCTTAT TATTATTAGA 7440
CAAGTGTGTG GTGTGAATAT GTCTTCAGAC TGGCACTTAG ACTTCCTATA AGTTCTTGCC 7500
TATCTAAGTT TTTCTAAATT GGGTTATTCT TGTAACATAT CTTAGATCTA GTACTCAACA 7560
CCACGTCACC ACCACAAAAG ATTTCTTATG CTCAAAAACA TATACATAGA AAGAACCTTC 7620
TAAACTACGA GAAACGTTTT GCTATGTAGT GTTATATGTC AACCACGTCT ATGAGAGTGC 7680
AAACGATAGG TTAATAAGTT TTCTCACTTG GCAATAAAAA TGATAAACAA ATATATTGTC 7740
TGATTAATTT ATTTTATATA GTTTTTTTAT AATTTCTTAT ATTAATTCGA ACTCATACAG 7800
CGCGTGAGAC TTTCTAGTTT AGTATAAAGT ACGTATTTTT GCAAAATCAA AATCGTAAAT 7860

ACATACATTT TAAAATGTTA AAAAAGATAA ATCCGTACAC CATTTAAAAA TGGCATTTTC 7920
CTAAGATTTT TTTCAAAAAA GGCATTTTAG ACAAGAACTA ATTACTACAA CTAAAATCTA 7980
CTAACTTTGG TTTTTATGTA TACATTTACG AGAGTCTACA CAAAAAAAAT ACATAAAAGA 8040
AGAAGTAGTA AATAATTAAA ACGTAAAAAA AAAGACTTTT CAAGAAGGCA GAAGAGTAGC 8100
ACTGTTGTGC GATTGTAAAA TCGTCTTGAT TGTTGTTTAT CCCACTGATA AGCCTACCCT 8160
TTTCAAAACT TGTTCTAAGT TTAAATTCTA TTTTTGAACA TGACATACAG TATAAGGCTT 8220
TTTAAAGATA TCATCTTGAT TTTGTTTCTT CCACAGGGAA GCCCTATCCT TTCTTACATA 8280
ATCTTTGTTA GATAATTTTT TATTATTTTC AAAAAAAATA AAATTGAACA TAAGTTTTCT 8340
CAAAGTAATA TGTTCTAACA ATAATAAACA TAATATCATT TTTTTGTTTT AAACTATAAA 8400
GGACTAACAT GGTAAAAAGT TGCAATATAT AAATGATAAT TTAAACTAAA AATTAGAATA 8460
TGGTAACTTT TTCTTCAACA ACATGCCACA TTCGGCTACA TGTCCACTAG GAAGTGTTAT 8520
TATAGAATCG TTAATGTTGG GTACGCTTAT GAAATTATCA ATGTTTGCTT AAATCTATGC 8580
TTAGAAAATT ACCAATATTA CCTTAAAACT ATATTTACGA ATGACCAATA TTGCTTAGAA 8640
CTATGCTTAT GAAATTACCA ATATTTTCTT AAAACTTAAA CACAAAACTC TTTAACAAAA 8700
AAAACTTTAT TTTTATTTTT ATTTTTTTGG CAAAAAAAAA AACTTTATTT ATAAAGTGAA 8760
AGTCTCCAGA TAATTTTGAA TTTCATTTTT CCAGTTTTTA TTTAGAATAA TTTTTCTTCA 8820
TTTACAAAAT AAAAGAAAAC CCTAGGGTTT AGGGTTTAGG GTTTAGGAAA AAGCGATGAT 8880
ATATTAATTG TTATGAAATG TTTTTTTAAA AATAGTTAAC CAAACATTTT TTTAAAGAGA 8940
GTTTAGTTTC ACAAGGCATT TGTAAATTAG AGTAATTATC AATAAAAATG GAAGACAATC 9000
TAATTATTAT TTAGCAAAAA CTATATTTAG GAAAATTAGT TAAAGTTTAG AAATATATCA 9060
TCATAGTGTC AAACTAATTA AAATTATTTA ATTTTGTGAT ATACGTGATC ATATAATTTT 9120
ATGAATATTT AATATTATGA TACATGTAAC TCAGTAAACC TAAATTTAGA AGAAAAGTCA 9180
AAATAATCAT AACCAATTTA GATTCAACTT CTACTTTTGT TCCAAGAAAA AAACACATGG 9240
TTTGTTTTGT GGGATACTAA TGACATCTAT CAAAATCTAT GAAACCAAAT CTAGA 9295

Sequence Number: 3
Sequence length: 18
Sequence type: Nucleic acid
Number of chains: One
Topology: Straight chain
Class of sequence: Other nucleic acid Synthetic DNA
Antisense: No
Sequence

TATCTAAAAA CGCAGTCG 18

Sequence Number: 4
Sequence length: 18
Sequence type: Nucleic acid
Number of chains: One
Topology: Straight chain
Class of sequence: Other nucleic acid Synthetic DNA

Antisense: Yes

Sequence

AAGATTCTCC TCCTAACG 18

Claims

Claim 1 DNA that codes for a protein that includes an amino acid sequence that possesses the action of controlling plant morphogenesis that possesses the amino acid sequence shown in Sequence Number 1 or alternatively that amino acid sequence in which one or a plurality of amino acid residues that do not affect the action of controlling plant morphogenesis are substituted, excised or inserted.

Claim 2 DNA that codes for antisense RNA that controls the expression of the DNA of Claim 1.

Claim 3 The DNA of Claim 2 further characterized by possessing a substantively complementary base sequence in at least one portion of the base sequence in Sequence Number 1.

Claim 4 Plants whose morphology has been transformed by the DNA of Claim 1.

Claim 5 Plants whose morphology has been transformed by the DNA of Claim 2 or Claim 3.

Claim 6 DNA that possesses at least portion of the sequences represented by Base Numbers 1 to 1752 in Sequence Number 2 and that controls the expression of the DNA of Claim 1.

Simplified description of the drawings

Figure 1 is a restriction enzyme map for the morphogene clone acquired through the practical embodiment.

Figure 2 is a drawing showing the T-DNA insertion site in the chromosomal DNA fragment inferred from the restriction enzyme map and Southern Analysis.

Figure 3 is a drawing showing the positions of the subclones of the sequences adjacent to the T-DNA insertion part

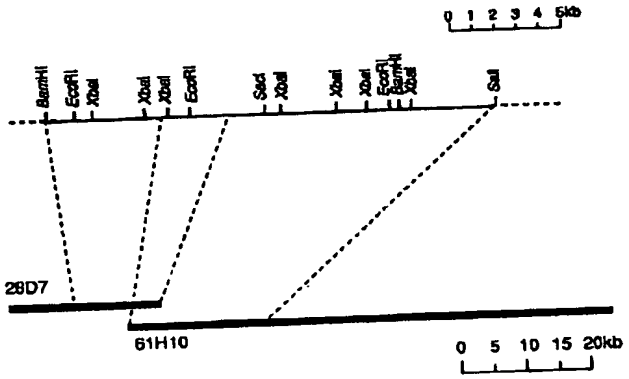


Figure 1

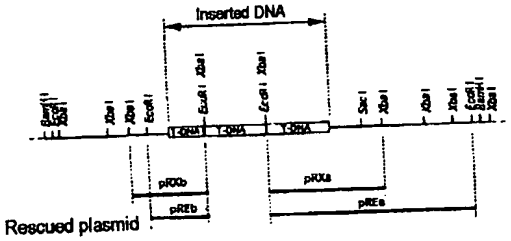


Figure 2

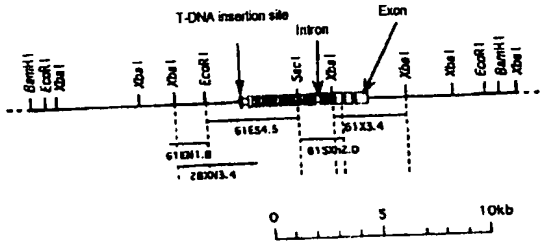


Figure 3

Abstract

Objective To obtain a gene that controls the morphogenesis of plants, and to employ such gene in order to control the morphology of plants

Means employed A gene that codes for a protein that includes an amino acid sequence that possesses the amino acid sequence shown in Sequence Number 1 that possesses the action of controlling plant morphogenesis or alternatively the sequence in which one or a plurality of amino acid residues that do not affect the action of controlling plant morphogenesis are substituted, deleted or inserted, or alternatively the transformation of the morphology of plants by means of DNA that expresses antisense RNA in such gene.

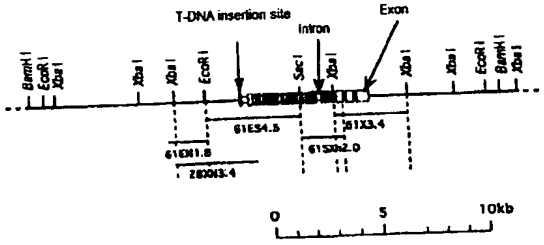


EXHIBIT D

(19)日本国特許庁（ＪＰ）　(12) **公開特許公報（A）**　(11)特許出願公開番号

特開平9－56382

(43)公開日　平成９年(1997)３月４日

(51)Int.Cl.⁶		識別記号	庁内整理番号	FI		技術表示箇所
C12N	15/09	ZNA	9162－4B	C12N	15/00	ZNAA
A01H	5/00	ZNA		A01H	5/00	ZNAA
C12N	5/10			C12N	5/00	C

審査請求　未請求　請求項の数６　ＯＬ　（全 17 頁）

(21)出願番号　特願平7－216187

(22)出願日　平成７年(1995)８月24日

(71)出願人　591178012
財団法人地球環境産業技術研究機構
京都府相楽郡木津町木津川台９丁目２番地

(71)出願人　591068458
株式会社三井農薬植物バイオ研究所
東京都港区赤坂２－５－27　八千代ビル４Ｆ

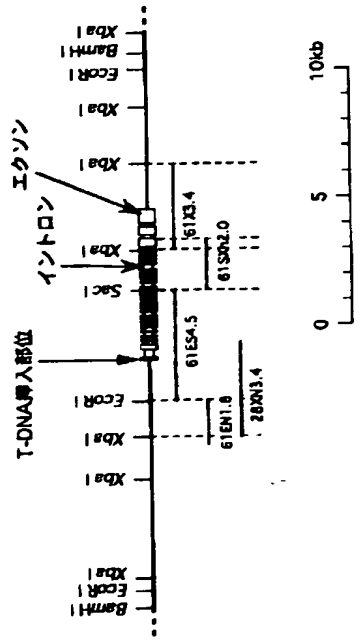
(72)発明者　光川　典宏
茨城県つくば市千現２－１－６　つくば研究支援センターＤ－６　株式会社三井農薬植物バイオ研究所内

(74)代理人　弁理士　遠山　勉　（外２名）

最終頁に続く

(54)【発明の名称】　植物の形態形成を制御するタンパク質をコードする遺伝子

(57)【要約】
【課題】　植物の形態形成を制御する遺伝子を取得し、これを用いて植物の形態を制御する。
【解決手段】　植物の形態形成を制御する活性を有し、配列番号１に示すアミノ酸配列又はこのアミノ酸配列において植物の形態形成を制御する活性に影響を与えない１又は２以上のアミノ酸残基の置換、欠失あるいは挿入を有するアミノ酸配列を含むタンパク質をコードする遺伝子、あるいはこの遺伝子に対するアンチセンスＲＮＡを発現するＤＮＡで植物を形質転換する。



Applicant: Josette Masle et al.
Serial No.: 10/519,135
Filed: August 15, 2005
Exhibit D

【特許請求の範囲】
【請求項１】　植物の形態形成を制御する活性を有し、配列番号１に示すアミノ酸配列又はこのアミノ酸配列において植物の形態形成を制御する活性に影響を与えない１又は２以上のアミノ酸残基の置換、欠失あるいは挿入を有するアミノ酸配列を含むタンパク質をコードするＤＮＡ。
【請求項２】　請求項１記載のＤＮＡの発現を抑制するアンチセンスＲＮＡをコードするＤＮＡ。
【請求項３】　配列番号１記載の塩基配列の少なくとも一部に実質的に相補的な塩基配列を有する請求項２記載のＤＮＡ。
【請求項４】　請求項１記載のＤＮＡで形質転換された植物体。
【請求項５】　請求項２または３記載のＤＮＡで形質転換された植物体。
【請求項６】　請求項１記載のＤＮＡの発現を制御するＤＮＡであって、配列番号２の塩基番号１～１７５２で表される配列の少なくとも一部を有するＤＮＡ。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】本発明は、植物の生態形成を制御する遺伝子ＤＮＡと、その遺伝子に対するアンチセンスＲＮＡをコードするＤＮＡ、並びにこれらのＤＮＡで形質転換された植物体に関し、植物の茎の長さ、および花序形態を調節する技術を提供するものである。
【０００２】
【従来の技術】植物形態は、遺伝的要因と環境要因によって影響されると考えられている。従来、茎の長さの短い植物や花序形態が変化した植物を作出する方法の一つとして、遺伝的にそれらの形態が異なる植物と交配することにより有用品種の形態を変える方法があるが、親品種よりも性状の優れた個体を安定して得ることは困難である。
【０００３】また、自然突然変異や突然変異誘発処理により形態に影響を与える遺伝子に変異が起こり、その遺伝子機能が低下することによって形態変化を示す個体を選抜することも可能であるが、有用遺伝形質を保持したまま、特定の形態形成を制御する遺伝子だけに変異が発生した個体を任意に作出することは非常に困難である。
【０００４】そこで植物の形態形成を制御する遺伝子を単離することができれば、該遺伝子に対するアンチセンスＲＮＡを発現するように組み込んだベクター（アンチセンスＲＮＡ発現ベクター）で植物を形質転換することにより、茎の長さが短くなった植物を作出することが可能になると考えられる。
【０００５】シロイヌナズナ（アラビドプシス・サリアナ（Arabidopsis thaliana））は、栄養期（vegetaive stage）から再生期（reproductive stage）への転移に花芽の伸長を伴い、蕾の生成及びそれに続く節間（internode）の伸長は強く連携しており、厳密に順序づけられた分枝様式（highly ordered branching pattern）を示す。シロイヌナズナの標準的なエコタイプ（ecotype）として知られているLandsberg erecta（ランズバーグ・エレクタ）株は、内因性エレクタ（erecta：わい化）変異を保持しており、異なる花芽構造を示す。頂部に花が密生するコンパクトな花芽を形成する。変異は、多面発現的（pleiotropic）であり、円形の葉と短く平坦な莢（silique）を有する（以上、Huang, I. et al., (1991) ＵＣＬＡキーストーンシンポジウムで発表された要旨）。
【０００６】また、Landsberg erecta株の変異と同一遺伝子座の変異を有することが遺伝学的に知られているわい化変異体が取得されており、er-101株、er-102株及びer-103株と命名されている。
【０００７】
【発明が解決しようとする課題】上述したように、植物のわい化と特定の遺伝子との関係は、ある程度経験的に知られているが、その遺伝子自体は未だ取得されていない。
【０００８】本発明は、上記観点からなされたものであり、植物の形態形成を制御する遺伝子を提供することを課題とする。また、該遺伝子又は該遺伝子に対するアンチセンスＲＮＡをコードするＤＮＡで植物を形質転換し、形態形成を制御する遺伝子の発現を促進あるいは抑制することによって、茎の長さが長くなった植物又は茎の長さが短くなった植物、あるいは花序形態の変化した植物を提供することも課題としている。
【０００９】
【課題を解決するための手段】上記目的を達成するために、本発明者らはシロイヌナズナの染色体ＤＮＡから植物の形態形成を制御する遺伝子をクローニングし、該遺伝子を組み込んだアンチセンス発現ベクターを得た後、該ベクターＤＮＡで形質転換した植物個体の形態が変化することを見いだし、本発明を完成するに至った。
【００１０】すなわち本発明は、植物の形態形成を制御する活性を有し、配列番号１に示すアミノ酸配列又はこのアミノ酸配列において植物の形態形成を制御する活性に影響を与えない１又は２以上のアミノ酸残基の置換、欠失あるいは挿入を有するアミノ酸配列を含むタンパク質をコードするＤＮＡである。
【００１１】また本発明は、上記のＤＮＡの発現を抑制するアンチセンスＲＮＡをコードするＤＮＡを提供する。このＤＮＡとしては、配列番号１記載の塩基配列の少なくとも一部に実質的に相補的な塩基配列を有するＤＮＡが挙げられる。
【００１２】本発明はさらに、前記タンパク質をコードするＤＮＡで形質転換された植物体、及び前記アンチセンスＲＮＡをコードするＤＮＡで形質転換された植物体を提供する。

【００１３】本発明のＤＮＡによりコードされるタンパク質は、形質の変化が著しい茎、花などでの発現量が多く、植物の形態形成の制御、特に茎の伸長に関わるタンパク質であり、このタンパク質をコードする遺伝子で植物を形質転換し、該遺伝子の発現量を増加させることによって、茎の伸長を促進できることが期待される。

【００１４】一方、前記アンチセンスＲＮＡをコードするＤＮＡ、すなわち植物の形態形成の制御に関する遺伝子の発現を抑制するＤＮＡで植物を形質転換することにより、形質転換植物の茎の伸長を抑制することができることが期待される。

【００１５】尚、本明細書において、「染色体ＤＮＡ」及び「染色体遺伝子」は、植物細胞の核染色体に含まれるＤＮＡ及びこのＤＮＡ上に存在する遺伝子をいう。また、本発明により提供される植物の形態形成の制御に関する遺伝子を、「形態制御遺伝子」または「本発明の遺伝子」ということがある。

【００１６】

【発明の実施の形態】以下、本発明の実施の形態を説明する。本発明の遺伝子は、例えば、形態形成の制御に関する変異を有する植物体から、その表現形質変異に関わる変異遺伝子を単離することにより取得することができる。さらに、得られた変異遺伝子の塩基配列に基づいて作製したオリゴヌクレオチドをプローブとするハイブリダイゼーション、あるいは変異遺伝子の塩基配列に基づいて作製した１対のオリゴヌクレオチドをプライマーとするＰＣＲ（ポリメラーゼ・チェイン・リアクション）により、野生型植物の染色体ＤＮＡから野生型遺伝子を取得することができる。

【００１７】形態形成の制御に関する変異を有する植物体の作製、この変異体からの本発明の遺伝子の単離法、及び本発明の遺伝子の利用法を詳細に説明する。尚、ＤＮＡの切断、連結、形質転換、遺伝子の塩基配列の決定、ハイブリダイゼーション等一般の遺伝子組換えに必要な方法は、各操作に使用する市販の酵素等に添付されている説明書や、Molecular cloning (Maniatis T. et al. Cold Spring HarborLaboratory Press)に記載されている。

【００１８】＜１＞植物の形態形成を制御する遺伝子の単離・同定

（１）形態形成の制御に関する変異を有する植物体の作製

植物、例えばシロイヌナズナの形態形成を制御する遺伝子に変異を起こさせるには、外来遺伝子を植物細胞に導入し、染色体ＤＮＡに挿入させることによって、挿入部位の遺伝子を破壊する変異誘発法（遺伝子破壊）を用いる。適用できる遺伝子導入法としては、アグロバクテリウムを用いる方法や、植物プロトプラスト細胞に対するエレクトロポーレーション法、ポリエチレングリコール法、マイクロインジェクション法などが挙げられる。これらの方法のうち、シロイヌナズナに対しては形質転換効率の高さと、遺伝子導入による変異以外の変異の誘起が少ない点から、アグロバクテリウムを用いる方法が有効である。

【００１９】ここでは、バイナリーベクター系（植物細胞にＤＮＡ導入可能なＴ－ＤＮＡ、大腸菌などの微生物で機能可能な複製起点、及び好ましくは植物細胞または微生物細胞の選択用のマーカー遺伝子を含むベクター系）を用いたアグロバクテリウム感染法によって外来遺伝子を植物細胞に導入し、形態形成を制御する遺伝子に関する変異植物を作製する方法を示す。

【００２０】シロイヌナズナ（Arabidopsis thaliana）の野生株に、Tiプラスミド由来のバイナリーベクター（マーカーとして、例えばハイグロマイシン耐性遺伝子及びアンピシリン耐性遺伝子を有する）を、インプランタ（in planta）アグロバクテリウム感染法（Chang S. S., Park S.K. et al. Plant J. 5, No.4(1994)）などにより感染させ、約６週間育成後に採種する。得られた種子をハイグロマイシンを含む寒天培地に播種し、生育させる。ハイグロマイシン耐性を示す形質転換植物をロックウールに移植して育成し、茎の伸長状態が野生型植物に比べて異なる変異体（erecta変異体）を視覚的に選択する。

【００２１】こうして得られる変異体は、植物の形態形成を制御する遺伝子にバイナリーベクターが挿入することによって変異している可能性が高い。

【００２２】（２）形態遺伝子の変異遺伝子の単離

上記のようにして得られる形態の変化した変異体から、その変異に関わる変異遺伝子を、いわゆるプラスミドレスキュー法を用いて単離する。すなわち、変異植物体からCell, 35 (1983) p.35記載の方法等で染色体ＤＮＡを調製して制限酵素で切断し、セルフライゲーションにより分子内の末端同士を連結する。得られた環状ＤＮＡが、染色体に挿入されたバイナリーベクターを含んでいれば、このＤＮＡ分子は大腸菌細胞で自律複製可能なプラスミドとして機能し、形質転換体はマーカー薬剤（例えばアンピシリン）耐性を示す。

【００２３】上記環状ＤＮＡで大腸菌を形質転換し、マーカー薬剤に耐性な形質転換体から組換えプラスミドＤＮＡを回収することによって、バイナリーベクターと共に形態制御遺伝子を含む染色体ＤＮＡ断片を得ることができる。

【００２４】あるいは、変異植物体から染色体ＤＮＡを調製し、適当な制限酵素で切断した後にプラスミドあるいはファージベクターに連結し、これで大腸菌を形質転換することにより染色体ライブラリーを作製する。このライブラリーからＴ－ＤＮＡ等をプローブとしてクローンを選抜することによっても、変異遺伝子断片を有するクローンを選択することができる。

【００２５】（３）形態制御遺伝子の野生型遺伝子の単

離
野生型のシロイヌナズナから、Ｐ１ファージベクター等を用いて染色体ＤＮＡライブラリーを作製し、上記のようにして得られる変異遺伝子断片をプローブとするハイブリダイゼーションによって、形態制御遺伝子の野生型遺伝子を単離することが出来る。また、変異遺伝子の塩基配列に基づいてプライマーを作製し、ＰＣＲにより野生型植物の染色体ＤＮＡから野生型遺伝子を増幅することによっても、本発明の遺伝子を取得することができる。

【００２６】後記実施例で得られたシロイヌナズナの形態制御遺伝子を含むＤＮＡ断片の塩基配列を、配列表配列番号２に示す。この遺伝子は、２７個のエクソン（exon）と２６個のイントロン（intron）を含んでいる。

【００２７】上記のように、本発明の遺伝子は多数のイントロンを含んでいる。エクソン部分、すなわち植物の形態形成を制御するタンパク質をコードするＤＮＡを得るには、形態制御遺伝子のｃＤＮＡを単離すればよい。ｃＤＮＡライブラリーは、シロイヌナズナの地上部組織からｍＲＮＡを抽出し、逆転写酵素を用いてｃＤＮＡを合成し、ポリメラーゼ反応によって２本鎖化したものをベクターに挿入し、大腸菌等を形質転換することにより作製することができる。ｃＤＮＡクローニングキットが市販されているのでこれらを使用してもよい。得られたｃＤＮＡライブラリーから、染色体遺伝子をプローブとして形態制御遺伝子ｃＤＮＡクローンを得る。

【００２８】上記のようにして後記実施例で得られたｃＤＮＡの塩基配列、及びこの塩基配列から推定されるアミノ酸配列を、配列表配列番号１に示す。この遺伝子の翻訳産物は、Nature、345（1990）p.743に記載されているRLK5と相同性を示した。このRLK5遺伝子は、細胞膜上に存在するレセプター様のプロテインキナーゼとして単離されたが、既知のプロテインキナーゼ遺伝子の塩基配列との相同性により単離されたもので、その遺伝子や翻訳産物の機能などは全く不明である。このRLK5遺伝子の発現パターンは、形態形成の制御に関する遺伝子とは異なり、地上部だけでなく、根でも発現しており機能的には両者は異なった遺伝子であると考えられる。

【００２９】＜２＞形態制御遺伝子の利用
本発明の遺伝子は、植物の形態形成の制御、特に茎の伸長に関わる遺伝子であり、この遺伝子で植物を形質転換し、該遺伝子の発現量を増加させることによって、茎の伸長を促進できることが期待される。

【００３０】本発明の遺伝子を用いて植物を形質転換するには、エレクトロポレーション（電気的穿孔法）あるいはアグロバクテリウムのＴｉプラスミドを利用する方法などの方法によって、プロトプラストにＤＮＡを導入すればよい。その際、本発明の遺伝子としては、染色体遺伝子を用いてもよいし、ｍＲＮＡから調製したｃＤＮＡを用いてもよい。

【００３１】一方、本発明の遺伝子の発現を抑制するアンチセンスＲＮＡ、すなわち形態制御遺伝子から転写されるｍＲＮＡの全長又はその少なくとも一部に相補的な配列を有するＲＮＡ、を発現するＤＮＡで植物を形質転換することにより、形質転換植物の茎の伸長を抑制することができることが期待される。

【００３２】アンチセンスＲＮＡを発現するＤＮＡは、アンチセンス鎖（センス鎖(コード鎖)に相補的な塩基配列を有する鎖）又は少なくともその一部をプロモーターの下流に連結することにより得られる。言い換えれば、コード鎖又はその少なくとも一部と相同な配列を含む２本鎖ＤＮＡを、本来の転写の向きと逆向きにしてプロモーターの下流に連結することにより得られる。尚、アンチセンス鎖も、染色体ＤＮＡあるいはｃＤＮＡから得られるが、染色体ＤＮＡを用いる場合にはイントロンは本発明遺伝子の発現を抑制する機能を果たすことが期待できないので、エクソン部分を用いることが好ましい。また、アンチセンス鎖の少なくとも一部としては、コード領域、５’非翻訳領域又は３’非翻訳領域のいずれも使用し得る。さらに、アンチセンスＲＮＡをコードするＤＮＡは、３’非翻訳領域に加え、ｍＲＮＡの３’末端に付加されるポリＡ鎖に相補的なポリｄＵをコードする配列を含んでいてもよい。

【００３３】本発明に使用し得るプロモーターとしては、ＣａＭＶ ３５Ｓプロモーター等が挙げられる。アンチセンスＲＮＡをコードするＤＮＡで植物を形質転換するで植物を形質転換するには、本発明の遺伝子で形質転換するのと同様にすればよい。

【００３４】さらに、本発明の遺伝子のコード領域の上流には、本発明の遺伝子の発現を制御する領域が含まれる。この領域としては、配列番号２の塩基番号１～１７５２で表される配列の少なくとも一部を含む領域、より具体的には塩基番号３９６～１７５２の領域が挙げられる。この領域は、植物細胞において遺伝子の発現制御に利用することができることが期待される。

【００３５】
【実施例】以下に、本発明を実施例によりさらに具体的に説明する。
（１）形態形成の制御に関する変異体植物の作製
シロイヌナズナ（アラビドプシス） エコタイプWS（Wassilewskija、LEHLE SEEDS 社より購入）株に、Tiプラスミド由来のバイナリーベクター（ハイグロマイシン耐性遺伝子及びアンピシリン耐性遺伝子を有し、大腸菌細胞内で自律複製可能）であるpGDW32を持つアグロバクテリウムEHA101株（Agrobacterium tumefacience）を、インプランタ（in planta）アグロバクテリウム感染法（Chang S.S.,Park S.K. et al. Plant J. 5, No.4(1994)）により感染させた。

【００３６】具体的には、以下のようにして行った。pGDW32を導入したアグロバクテリウムを、抗生物質（ハイ

グロマイシン）を含むＬＢ液体培地で一晩増殖させた。この培養液１０μｌをＧａｍｂｏｒｇ’ｓ Ｂ５培地（ショ糖２％を含み、ｐＨ５．５）１９０μｌに希釈（1/20）した。播種から１９～２３日後、抽台を始めたばかりの生育段階でアグロバクテリウム処理を行った。

【００３７】伸び始めたばかりの花茎をその基部において１１号メスを用いて切り取り、２６Ｇ×１／２サイズの注射針を用いて切り口からロゼット茎の中央部を貫通させた。この傷口にアグロバクテリウム希釈液１μｌを注入した。この間、照度を３０００～４０００ルクスにした。接種３日後に植物をロックファイバーのミニポットに移植し、その後は通常の栽培を行った。接種から約６週間後、早く形成された約半分の莢から種子を収穫した。

【００３８】得られた種子を、10μg/mlのハイグロマイシンを含むB5寒天培地に播種し、温度22℃、湿度30-40%、照度6000-12000ルクスの照明下で、明期12時間/暗期12時間の光周期で生育させた。栄養素は1/1000に希釈したHYPONeX（村上物産（株）製）を用いた。ハイグロマイシン耐性を示すpGDW32のT-DNA挿入配列を持つ形質転換植物をロックウールに移植して育成し、T4種子（第４世代の種子）を採取した。この際、視覚的に検索を行い、茎の伸長状態が野生型植物に比べて異なる変異体を得た。

【００３９】（２）変異遺伝子の単離

アラビドプシスのゲノムDNAを、Cell, 35 (1983) p.35に記載の方法等により調製した。

【００４０】-80℃で凍結したアラビドプシス組織（根） 5g を、乳鉢を用いて液体窒素中で微粉末になるまで粉砕し、これを25ml の DNA 単離用バッファー（50 mM Tris-HCl,pH7.5, 0.2M NaCl, 20mM EDTA-Na2, 2% N-ラウロイルザルコシンナトリウム塩, 3g/ml 尿素, 5% TE飽和フェノール）を加えて撹拌し、25ml のフェノール/クロロホルムを加えた後、1.5ml の10% SDS（ドデシル硫酸ナトリウム） を加えて10分間、室温で緩やかに撹拌した。これを6000prm で10分遠心した後、水層に25ml のフェノール/クロロホルムを加えた後、再度6000prm で10分遠心した。この水層に15mlのエタノールを加え、撹拌した後直ちに、6000prm で10分遠心した。上清を捨て、沈殿に25mlの70%エタノールを加えて、ボルテックスにより撹拌し、6000prm で10分遠心した後上清を捨て、減圧下で乾燥した。これを400μlのTE,pH8.0（10μg/μl RNase）に溶解した。

【００４１】上記のようにして、形態形成の制御に関する変異体から調製した 200μg の染色体DNA を、続いて CsCl 超遠心法により精製した。このうち、１μg の染色体 DNA を EcoRI又はXbaIで切断し、フェノール抽出、エタノール沈殿で制限酵素断片を精製した後、セルフライゲーションにより分子内の末端同士を連結し、得られた環状ＤＮＡで大腸菌（XL1-Blue MRF'（STRATAGENE社から購入））を形質転換した。その結果、約1000コロニーのアンピシリン耐性を示す形質転換株を得ることができた。得られた耐性コロニーからプラスミドＤＮＡを抽出して解析を行った。

【００４２】上記のようにしてレスキューされたプラスミドのうち、EcoRIで切断したDNAから得られたものをpREa及びpREb、XbaIで切断したDNAから得られたものをpRXa及びpRXbと命名した。

【００４３】（３）変異遺伝子の単離

上記のようにしてレスキューされたプラスミドに含まれている染色体ＤＮＡ断片をプローブとして、染色体ＤＮＡライブラリーからの変異遺伝子の単離を行った。Ｐ１ファージベクター（DU PONT社から購入した）を用いて、The Plant Journal, 7 (1995) p.351に記載の方法にしたがって、アラビドプシス エコタイプ コロンビア株の核ＤＮＡライブラリーを調製した。一方、上記pRXbとpREaからそれぞれ調製したEcoRI-XbaI断片を^{32}Ｐ標識し、これらをプローブとして、プラークハイブリダイゼーションを行った。その結果、２つの陽性クローンが得られ、それぞれ28D7、61H10と命名した。これらのクローンの制限酵素地図の作製を行い、28D7は約25kb、61H10は約75kbのアラビドプシス染色体ＤＮＡ由来の挿入断片を持つことがわかった（図１）。

【００４４】さらに、塩基配列決定のために挿入ＤＮＡ断片に含まれる変異遺伝子のサブクローニングを実施した。この際、T-DNAに隣接する配列の制限酵素地図と、挿入断片のサザン解析を実施し、挿入断片上のT-DNA挿入部位を推定した（図２）。この挿入部位の近傍の染色体ＤＮＡ配列をBluescript II SK+にサブクローニングした。得られたサブクローンの挿入断片上の位置を図３に示す。

【００４５】（４）形態形成の制御に関する遺伝子のcDNAクローンの単離

アラビドプシス エコタイプ コロンビア株の地上部組織からｍＲＮＡを調製し、Molecular cloning（Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989)、ALaboraroty Manual, second edition, Cold Spring Harbor LaboratoryPress, Cold Spring Harbor, NY）の方法によりｃＤＮＡを調製し、λYES（CLONTECH社から購入した）をベクターとしてｃＤＮＡライブラリーを作製した。一方、前記pRXbとpREaからT-DNAに隣接する配列を切り出し、これを^{32}Ｐ標識した。これらをプローブとして、ファージライブラリーの300,000プラークを対象としてプラークハイブリダイゼーションを行った。こうして選択したプラスミドをｐＫＵＴ１６１とした。

【００４６】（５）形態形成の制御に関する遺伝子の解析

ｐＫＵＴ１６１に含まれるｃＤＮＡクローンと染色体遺伝子の塩基配列を決定した。ｃＤＮＡの塩基配列及びこの配列から推定されるアミノ酸配列を配列表配列番号１

に示す。また、染色体遺伝子を含むＤＮＡ断片の塩基配列を配列番号２に示す。

【００４７】cDNAの塩基配列から推定されるアミノ酸配列の解析を行ったところ、この遺伝子の翻訳産物は、Nature、345（1990）p.743に記載されているRLK5と相同性を示した。このRLK5遺伝子は、細胞膜上に存在するレセプター様のプロテインキナーゼとして単離されたが、既知のプロテインキナーゼ遺伝子の塩基配列との相同性により単離されたもので、その遺伝子や翻訳産物の機能などは全く不明である。このRLK5遺伝子の発現パターンは、形態形成の制御に関する遺伝子とは異なり、地上部だけでなく、根でも発現しており機能的には両者は異なった遺伝子であると考えられる。

【００４８】（６）形態形成の制御に関する変異体の解析

前記（１）で得られた形態形成の制御に関する変異体は、茎の伸長状態から、Landsberg erecta株のエレクタ変異と同一遺伝子座の変異であると考えられた。そこでまず、先のLandsberg erecta株、er-103株と本発明で単離した変異体の遺伝的な相補試験を行った。その結果これらの変異はすべて同一遺伝子座の変異に起因することが判明し、単離した変異体をer-104株とした。

【００４９】Landsberg erecta株、er-103株からｍＲＮＡを単離し、ＲＴ－ＰＣＲ（逆転写－ＰＣＲ）法で増幅した。すなわち、ｍＲＮＡを鋳型とし、配列番号４に示す塩基配列を有するオリゴヌクレオチドをプライマーとして逆転写反応を行い、続いて得られたｃＤＮＡを鋳型とし、前記オリゴヌクレオチドを３’側プライマー及び配列番号３に示す塩基配列を有するオリゴヌクレオチドを５’側プライマーに用いてＰＣＲ反応を行った。

【００５０】得られた増幅産物を鋳型とし、上記オリゴヌクレオチドプライマーを用いたダイレクトシーケンス法により、塩基配列を決定した。その結果、Landsberg erecta株では配列番号１の塩基番号2299番のＴがＡに置換したのに伴い、アミノ酸番号750番のイソロイシン残基がリジン残基に置換していた。er-103株では配列表配列番号１の塩基番号896番のＧがＡに置換することにより、アミノ酸番号282番のメチオニン残基がイソロイシン残基に置換していた。

【００５１】これら全ての変異体間で、形態形成の制御に関する遺伝子の塩基配列の一部が、アミノ酸配列の変化を引き起こす変異を持つことが判明し、この遺伝子が形態形成を制御する遺伝子（形態制御遺伝子）であることが同定できた。

【００５２】（７）形態制御遺伝子の発現解析

上記のようにして単離されたcDNA、並びに染色体遺伝子が形態形成を制御する遺伝子であることを確認するもう一つの方法として、変異体植物の各組織での発現量分析、及び野生型植物と変異体植物の間での遺伝子の発現量の変化を、ノーザン分析により比較分析した。プローブにはｃＤＮＡクローンを用いた。Landsberg erecta変異体と野生型植物の個体全体を用いたノーザン分析の結果から、形態制御遺伝子の発現（mRNAの生成）が、野生型植物では強く観察されるのに対して、変異体ではほとんど認められなかった。

【００５３】また、野生型植物の各組織での形態制御遺伝子の発現量分析を行った結果、変異体で形質の変化が著しい茎、花などでの発現量が多く、このことからもこの遺伝子が茎の伸長に関わる遺伝子であることが確認された。

【００５４】さらに、上述したように、形態制御遺伝子の発現は部位及び時期特異的であるので、本発明の遺伝子の発現制御領域は、植物における外来遺伝子発現の制御に利用し得る。そのような発現制御領域としては、配列番号２の塩基番号１～１７５２で表される配列の少なくとも一部を含む領域、より具体的には塩基番号３９６～１７５２の領域が挙げられる。

【００５５】

【発明の効果】本発明により、植物の形態形成を制御する遺伝子が提供される。この遺伝子の発現量を増やすことによって、茎の伸長を促進できることが期待される。また、この遺伝子に対するアンチセンスＲＮＡを発現するＤＮＡ配列で植物を形質転換することにより、茎の伸長を抑制できることが期待される。

【００５６】

【配列表】

配列番号：１
配列の長さ：3176
配列の型：核酸
鎖の数：二本鎖
トポロジー：直鎖状
配列の種類：cDNA to mRNA
起源
生物名：シロイヌナズナ（Arabidopsis thaliana）
株名：コロンビア
配列の特徴：
特徴を表す記号：　CDS
存在位置：　51..2978

```
CTTTTAAAGT ATATCTAAAA ACGCAGTCGT TTTAAGACTG TGTGTGAGAA ATG GCT        56
                                                        Met Ala
                                                          1

CTG TTT AGA GAT ATT GTT CTT CTT GGG TTT CTC TTC TGC TTG AGC TTA       104
Leu Phe Arg Asp Ile Val Leu Leu Gly Phe Leu Phe Cys Leu Ser Leu
  5                  10                  15
```

```
GTA GCT ACT GTG ACT TCA GAG GAG GGA GCA ACG TTG CTG GAG ATT AAG      152
Val Ala Thr Val Thr Ser Glu Glu Gly Ala Thr Leu Leu Glu Ile Lys
     20                  25                  30
AAG TCA TTC AAA GAT GTG AAC AAT GTT CTT TAT GAC TGG ACA ACT TCA      200
Lys Ser Phe Lys Asp Val Asn Asn Val Leu Tyr Asp Trp Thr Thr Ser
 35                  40                  45                  50
CCT TCT TCG GAT TAT TGT GTC TGG AGA GGT GTG TCT TGT GAA AAT GTC      248
Pro Ser Ser Asp Tyr Cys Val Trp Arg Gly Val Ser Cys Glu Asn Val
                 55                  60                  65
ACC TTC AAT GTT GTT GCT CTT AAT TTG TCA GAT TTG AAT CTT GAT GGA      296
Thr Phe Asn Val Val Ala Leu Asn Leu Ser Asp Leu Asn Leu Asp Gly
             70                  75                  80
GAA ATC TCA CCT GCT ATT GGA GAT CTC AAG AGT CTC TTG TCA ATT GAT      344
Glu Ile Ser Pro Ala Ile Gly Asp Leu Lys Ser Leu Leu Ser Ile Asp
         85                  90                  95
CTG CGA GGT AAT CGC TTG TCT GGA CAA ATC CCT GAT GAG ATT GGT GAC      392
Leu Arg Gly Asn Arg Leu Ser Gly Gln Ile Pro Asp Glu Ile Gly Asp
    100                 105                 110
TGT TCT TCT TTG CAA AAC TTA GAC TTA TCC TTC AAT GAA TTA AGT GGT      440
Cys Ser Ser Leu Gln Asn Leu Asp Leu Ser Phe Asn Glu Leu Ser Gly
115                 120                 125                 130
GAC ATA CCG TTT TCG ATT TCG AAG TTG AAG CAA CTT GAG CAG CTG ATT      488
Asp Ile Pro Phe Ser Ile Ser Lys Leu Lys Gln Leu Glu Gln Leu Ile
                135                 140                 145
CTG AAG AAT AAC CAA TTG ATA GGA CCG ATC CCT TCA ACA CTT TCA CAG      536
Leu Lys Asn Asn Gln Leu Ile Gly Pro Ile Pro Ser Thr Leu Ser Gln
            150                 155                 160
ATT CCA AAC CTG AAA ATT CTG GAC TTG GCA CAG AAT AAA CTC AGT GGT      584
Ile Pro Asn Leu Lys Ile Leu Asp Leu Ala Gln Asn Lys Leu Ser Gly
        165                 170                 175
GAG ATA CCA AGA CTT ATT TAC TGG AAT GAA GTT CTT CAG TAT CTT GGG      632
Glu Ile Pro Arg Leu Ile Tyr Trp Asn Glu Val Leu Gln Tyr Leu Gly
    180                 185                 190
TTG CGA GGA AAC AAC TTA GTC GGT AAC ATT TCT CCA GAT TTG TGT CAA      680
Leu Arg Gly Asn Asn Leu Val Gly Asn Ile Ser Pro Asp Leu Cys Gln
195                 200                 205                 210
CTG ACT GGT CTT TGG TAT TTT GAC GTA AGA AAC AAC AGT TTG ACT GGT      728
Leu Thr Gly Leu Trp Tyr Phe Asp Val Arg Asn Asn Ser Leu Thr Gly
                215                 220                 225
AGT ATA CCT GAG ACG ATA GGA AAT TGC ACT GCC TTC CAG GTT TTG GAC      776
Ser Ile Pro Glu Thr Ile Gly Asn Cys Thr Ala Phe Gln Val Leu Asp
            230                 235                 240
TTG TCC TAC AAT CAG CTA ACT GGT GAG ATC CCT TTT GAC ATC GGC TTC      824
Leu Ser Tyr Asn Gln Leu Thr Gly Glu Ile Pro Phe Asp Ile Gly Phe
        245                 250                 255
CTG CAA GTT GCA ACA TTA TCA TTG CAA GGC AAT CAA CTC TCT GGG AAG      872
Leu Gln Val Ala Thr Leu Ser Leu Gln Gly Asn Gln Leu Ser Gly Lys
    260                 265                 270
ATT CCA TCA GTG ATT GGT CTC ATG CAA GCC CTT GCA GTC TTA GAT CTA      920
Ile Pro Ser Val Ile Gly Leu Met Gln Ala Leu Ala Val Leu Asp Leu
```

```
275                 280                 285                 290
AGT GGC AAC TTG TTG AGT GGA TCT ATT CCT CCG ATT CTC GGA AAT CTT      968
Ser Gly Asn Leu Leu Ser Gly Ser Ile Pro Pro Ile Leu Gly Asn Leu
                295                 300                 305
ACT TTC ACC GAG AAA TTG TAT TTG CAC AGT AAC AAG CTG ACT GGT TCA     1016
Thr Phe Thr Glu Lys Leu Tyr Leu His Ser Asn Lys Leu Thr Gly Ser
            310                 315                 320
ATT CCA CCT GAG CTT GGA AAC ATG TCA AAA CTC CAT TAC CTG GAA CTC     1064
Ile Pro Pro Glu Leu Gly Asn Met Ser Lys Leu His Tyr Leu Glu Leu
        325                 330                 335
AAT GAT AAT CAT CTC ACG GGT CAT ATA CCA CCA GAG CTT GGG AAG CTT     1112
Asn Asp Asn His Leu Thr Gly His Ile Pro Pro Glu Leu Gly Lys Leu
    340                 345                 350
ACT GAC TTG TTT GAT CTG AAT GTG GCC AAC AAT GAT CTG GAA GGA CCT     1160
Thr Asp Leu Phe Asp Leu Asn Val Ala Asn Asn Asp Leu Glu Gly Pro
355                 360                 365                 370
ATA CCT GAT CAT CTG AGC TCT TGC ACA AAT CTA AAC AGC TTA AAT GTT     1208
Ile Pro Asp His Leu Ser Ser Cys Thr Asn Leu Asn Ser Leu Asn Val
                375                 380                 385
CAT GGG AAC AAG TTT AGT GGC ACT ATA CCC CGA GCA TTT CAA AAG CTA     1256
His Gly Asn Lys Phe Ser Gly Thr Ile Pro Arg Ala Phe Gln Lys Leu
            390                 395                 400
GAA AGT ATG ACT TAC CTT AAT CTG TCC AGC AAC AAT ATC AAA GGT CCA     1304
Glu Ser Met Thr Tyr Leu Asn Leu Ser Ser Asn Asn Ile Lys Gly Pro
        405                 410                 415
ATC CCG GTT GAG CTA TCT CGT ATC GGT AAC TTA GAT ACA TTG GAT CTT     1352
Ile Pro Val Glu Leu Ser Arg Ile Gly Asn Leu Asp Thr Leu Asp Leu
    420                 425                 430
TCC AAC AAC AAG ATA AAT GGA ATC ATT CCT TCT TCC CTT GGT GAT TTG     1400
Ser Asn Asn Lys Ile Asn Gly Ile Ile Pro Ser Ser Leu Gly Asp Leu
435                 440                 445                 450
GAG CAT CTT CTC AAG ATG AAC TTG AGT AGA AAT CAT ATA ACT GGT GTA     1448
Glu His Leu Leu Lys Met Asn Leu Ser Arg Asn His Ile Thr Gly Val
                455                 460                 465
GTT CCA GGC GAC TTT GGA AAT CTA AGA AGC ATC ATG GAA ATA GAT CTT     1496
Val Pro Gly Asp Phe Gly Asn Leu Arg Ser Ile Met Glu Ile Asp Leu
            470                 475                 480
TCA AAT AAT GAT ATC TCT GGC CCA ATT CCA GAA GAG CTT AAC CAA TTA     1544
Ser Asn Asn Asp Ile Ser Gly Pro Ile Pro Glu Glu Leu Asn Gln Leu
        485                 490                 495
CAG AAC ATA ATT TTG CTG AGA CTG GAA AAT AAT AAC CTG ACT GGT AAT     1592
Gln Asn Ile Ile Leu Leu Arg Leu Glu Asn Asn Asn Leu Thr Gly Asn
    500                 505                 510
GTT GGT TCA TTA GCC AAC TGT CTC AGT CTC ACT GTA TTG AAT GTA TCT     1640
Val Gly Ser Leu Ala Asn Cys Leu Ser Leu Thr Val Leu Asn Val Ser
515                 520                 525                 530
CAT AAC AAC CTC GTA GGT GAT ATC CCT AAG AAC AAT AAC TTC TCA AGA     1688
His Asn Asn Leu Val Gly Asp Ile Pro Lys Asn Asn Asn Phe Ser Arg
                535                 540                 545
TTT TCA CCA GAC AGC TTC ATT GGC AAT CCT GGT CTT TGC GGT AGT TGG     1736
```

```
Phe Ser Pro Asp Ser Phe Ile Gly Asn Pro Gly Leu Cys Gly Ser Trp
            550                 555                 560
CTA AAC TCA CCG TGT CAT GAT TCT CGT CGA ACT GTA CGA GTG TCA ATC    1784
Leu Asn Ser Pro Cys His Asp Ser Arg Arg Thr Val Arg Val Ser Ile
        565                 570                 575
TCT AGA GCA GCT ATT CTT GGA ATA GCT ATT GGG GGA CTT GTG ATC CTT    1832
Ser Arg Ala Ala Ile Leu Gly Ile Ala Ile Gly Gly Leu Val Ile Leu
    580                 585                 590
CTC ATG GTC TTA ATA GCA GCT TGC CGA CCG CAT AAT CCT CCT CCT TTT    1880
Leu Met Val Leu Ile Ala Ala Cys Arg Pro His Asn Pro Pro Pro Phe
595                 600                 605                 610
CTT GAT GGA TCA CTT GAC AAA CCA GTA ACT TAT TCG ACA CCG AAG CTC    1928
Leu Asp Gly Ser Leu Asp Lys Pro Val Thr Tyr Ser Thr Pro Lys Leu
                615                 620                 625
GTC ATC CTT CAT ATG AAC ATG GCA CTC CAC GTT TAC GAG GAT ATC ATG    1976
Val Ile Leu His Met Asn Met Ala Leu His Val Tyr Glu Asp Ile Met
            630                 635                 640
AGA ATG ACA GAG AAT CTA AGT GAG AAG TAT ATC ATT GGG CAC GGA GCA    2024
Arg Met Thr Glu Asn Leu Ser Glu Lys Tyr Ile Ile Gly His Gly Ala
        645                 650                 655
TCA AGC ACT GTA TAC AAA TGT GTT TTG AAG AAT TGT AAA CCG GTT GCG    2072
Ser Ser Thr Val Tyr Lys Cys Val Leu Lys Asn Cys Lys Pro Val Ala
    660                 665                 670
ATT AAG CGG CTT TAC TCT CAC AAC CCA CAG TCA ATG AAA CAG TTT GAA    2120
Ile Lys Arg Leu Tyr Ser His Asn Pro Gln Ser Met Lys Gln Phe Glu
675                 680                 685                 690
ACA GAA CTC GAG ATG CTA AGT AGC ATC AAG CAC AGA AAT CTT GTG AGC    2168
Thr Glu Leu Glu Met Leu Ser Ser Ile Lys His Arg Asn Leu Val Ser
                695                 700                 705
CTA CAA GCT TAT TCC CTC TCT CAC TTG GGG AGT CTT CTG TTC TAT GAC    2216
Leu Gln Ala Tyr Ser Leu Ser His Leu Gly Ser Leu Leu Phe Tyr Asp
            710                 715                 720
TAT TTG GAA AAT GGT AGC CTC TGG GAT CTT CTT CAT GGC CCT ACG AAG    2264
Tyr Leu Glu Asn Gly Ser Leu Trp Asp Leu Leu His Gly Pro Thr Lys
        725                 730                 735
AAA AAG ACT CTT GAT TGG GAC ACA CGG CTT AAG ATA GCA TAT GGT GCA    2312
Lys Lys Thr Leu Asp Trp Asp Thr Arg Leu Lys Ile Ala Tyr Gly Ala
    740                 745                 750
GCA CAA GGT TTA GCT TAT CTA CAC CAT GAC TGT AGT CCA AGG ATC ATT    2360
Ala Gln Gly Leu Ala Tyr Leu His His Asp Cys Ser Pro Arg Ile Ile
755                 760                 765                 770
CAC AGA GAC GTG AAG TCG TCC AAC ATT CTC TTG GAC AAA GAC TTA GAG    2408
His Arg Asp Val Lys Ser Ser Asn Ile Leu Leu Asp Lys Asp Leu Glu
                775                 780                 785
GCT CGT TTG ACA GAT TTT GGA ATA GCG AAA AGC TTG TGT GTG TCA AAG    2456
Ala Arg Leu Thr Asp Phe Gly Ile Ala Lys Ser Leu Cys Val Ser Lys
            790                 795                 800
TCA CAT ACT TCA ACT TAC GTG ATG GGC ACG ATA GGT TAC ATA GAC CCC    2504
Ser His Thr Ser Thr Tyr Val Met Gly Thr Ile Gly Tyr Ile Asp Pro
        805                 810                 815
```

```
GAG TAT GCT CGC ACT TCA CGG CTC ACT GAG AAA TCC GAT GTC TAC AGT     2552
Glu Tyr Ala Arg Thr Ser Arg Leu Thr Glu Lys Ser Asp Val Tyr Ser
    820                 825                 830
TAT GGA ATA GTC CTT CTT GAG CTG TTA ACC CGA AGG AAA GCC GTT GAT     2600
Tyr Gly Ile Val Leu Leu Glu Leu Leu Thr Arg Arg Lys Ala Val Asp
835                 840                 845                 850
GAC GAA TCC AAT CTC CAC CAT CTG ATA ATG TCA AAG ACG GGG AAC AAT     2648
Asp Glu Ser Asn Leu His His Leu Ile Met Ser Lys Thr Gly Asn Asn
                855                 860                 865
GAA GTG ATG GAA ATG GCA GAT CCA GAC ATC ACA TCG ACG TGT AAA GAT     2696
Glu Val Met Glu Met Ala Asp Pro Asp Ile Thr Ser Thr Cys Lys Asp
            870                 875                 880
CTC GGT GTG GTG AAG AAA GTT TTC CAA CTG GCA CTC CTA TGC ACC AAA     2744
Leu Gly Val Val Lys Lys Val Phe Gln Leu Ala Leu Leu Cys Thr Lys
        885                 890                 895
AGA CAG CCG AAT GAT CGA CCC ACA ATG CAC CAG GTG ACT CGT GTT CTC     2792
Arg Gln Pro Asn Asp Arg Pro Thr Met His Gln Val Thr Arg Val Leu
    900                 905                 910
GGC AGT TTT ATG CTA TCG GAA CAA CCA CCT GCT GCG ACT GAC ACG TCA     2840
Gly Ser Phe Met Leu Ser Glu Gln Pro Pro Ala Ala Thr Asp Thr Ser
915                 920                 925                 930
GCG ACG CTG GCT GGT TCG TGC TAC GTC GAT GAG TAT GCA AAT CTC AAG     2888
Ala Thr Leu Ala Gly Ser Cys Tyr Val Asp Glu Tyr Ala Asn Leu Lys
                935                 940                 945
ACT CCT CAT TCT GTC AAT TGC TCT TCC ATG AGT GCT TCT GAT GCT CAA     2936
Thr Pro His Ser Val Asn Cys Ser Ser Met Ser Ala Ser Asp Ala Gln
            950                 955                 960
CTG TTT CTT CGG TTT GGA CAA GTT ATT TCT CAG AAC AGT GAG             2978
Leu Phe Leu Arg Phe Gly Gln Val Ile Ser Gln Asn Ser Glu
        965                 970                 975
TAGTTTTTCG TTAGGAGGAG AATCTTTAAA ACGGTATCTT TTCGTTGCGT TAAGCTGTTA  3038
GAAAAATTAA TGTCTCATGT AAAGTATTAT GCACTGCCTT ATTATTATTA GACAAGTGTG  3098
TGGTGTGAAT ATGTCTTCAG ACTGGCACTT AGACTTCCAA AAAAAAAAAA AAAAAAAAAA  3158
AAAAAAAAAA AAAAAAAA                                               3176
```

【００５７】配列番号：２

配列の長さ：9295
配列の型：核酸
鎖の数：二本鎖
トポロジー：直鎖状
配列の種類：genomic DNA
起源
生物名：シロイヌナズナ（Arabidopsis thaliana）
株名：コロンビア
配列の特徴：
特徴を表す記号： exon
存在位置： 1803..1881
配列の特徴：
特徴を表す記号： intron
存在位置： 1882..2227
配列の特徴：
特徴を表す記号： exon
存在位置： 2228..2366
配列の特徴：
特徴を表す記号： intron
存在位置： 2367..2467
配列の特徴：
特徴を表す記号： intron
存在位置： 2540..2643
配列の特徴：
特徴を表す記号： exon
存在位置： 2468..2539
配列の特徴：
特徴を表す記号： exon
存在位置： 2644..2715
配列の特徴：
特徴を表す記号： intron

存在位置： 2716..2809
配列の特徴：
特徴を表す記号： exon
存在位置： 2810..2878
配列の特徴：
特徴を表す記号： intron
存在位置： 2879..2968
配列の特徴：
特徴を表す記号： exon
存在位置： 2969..3040
配列の特徴：
特徴を表す記号： intron
存在位置： 3041..3118
配列の特徴：
特徴を表す記号： exon
存在位置： 3119..3190
配列の特徴：
特徴を表す記号： intron
存在位置： 3191..3266
配列の特徴：
特徴を表す記号： exon
存在位置： 3267..3338
配列の特徴：
特徴を表す記号： intron
存在位置： 3339..3421
配列の特徴：
特徴を表す記号： exon
存在位置： 3422..3493
配列の特徴：
特徴を表す記号： intron
存在位置： 3494..3586
配列の特徴：
特徴を表す記号： exon
存在位置： 3587..3655
配列の特徴：
特徴を表す記号： intron
存在位置： 3656..3740
配列の特徴：
特徴を表す記号： exon
存在位置： 3741..3812
配列の特徴：
特徴を表す記号： intron
存在位置： 3813..3888
配列の特徴：
特徴を表す記号： exon
存在位置： 3889..3960
配列の特徴：
特徴を表す記号： intron
存在位置： 3961..4048
配列の特徴：
特徴を表す記号： exon
存在位置： 4049..4120
配列の特徴：
特徴を表す記号： intron
存在位置： 4121..4209
配列の特徴：
特徴を表す記号： exon
存在位置： 4210..4281
配列の特徴：
特徴を表す記号： intron
存在位置： 4282..4349
配列の特徴：
特徴を表す記号： exon
存在位置： 4350..4421
配列の特徴：
特徴を表す記号： intron
存在位置： 4422..4508
配列の特徴：
特徴を表す記号： exon
存在位置： 4509..4580
配列の特徴：
特徴を表す記号： intron
存在位置： 4581..4706
配列の特徴：
特徴を表す記号： exon
存在位置： 4707..4778
配列の特徴：
特徴を表す記号： intron
存在位置： 4779..4860
配列の特徴：
特徴を表す記号： exon
存在位置： 4861..4932
配列の特徴：
特徴を表す記号： intron
存在位置： 4933..5018
配列の特徴：
特徴を表す記号： exon
存在位置： 5019..5090
配列の特徴：
特徴を表す記号： intron
存在位置： 5091..5176
配列の特徴：
特徴を表す記号： exon
存在位置： 5177..5248
配列の特徴：
特徴を表す記号： intron
存在位置： 5249..5412
配列の特徴：
特徴を表す記号： exon
存在位置： 5413..5481

配列の特徴：
特徴を表す記号： intron
存在位置： 5482..5576
配列の特徴：
特徴を表す記号： exon
存在位置： 5577..5648
配列の特徴：
特徴を表す記号： intron
存在位置： 5649..5726
配列の特徴：
特徴を表す記号： exon
存在位置： 5727..5800
配列の特徴：
特徴を表す記号： intron
存在位置： 5801..5882
配列の特徴：
特徴を表す記号： exon
存在位置： 5883..6011
配列の特徴：
特徴を表す記号： exon
存在位置： 6096..6443
配列の特徴：
特徴を表す記号： intron
存在位置： 6012..6095
配列の特徴：
特徴を表す記号： intron
存在位置： 6444..6519
配列の特徴：
特徴を表す記号： exon
存在位置： 6520..6890
配列の特徴：
特徴を表す記号： intron
存在位置： 6891..6974
配列の特徴：
特徴を表す記号： exon
存在位置： 6975..7328

配列

GAATTCAAAG GAATAAGCAT CGGAGACGAT TTAATGTTAC CTCTTGACGT ATTTATCCAA 60
TTTATCCATT AAGCCACCAG CCATAGCATC TGATCATCAT CATCAACATA TAAATAACCA 120
AATTTGAAAT GAACAAAAGT CGAATTGGTG ATATTGAAAA TCGAGTTCGT GAAATTGAGA 180
ATCGGATTGG TGAATTTGAA GAGAGATGCG TGTACCGTTA GGGAGGAGGA GGAGACGGGA 240
GAGAAAAAAG GAGACGGAGA TAACTCGCCG GCTCTGTTTC CATGGCGGAG GTGATAATGT 300
AGCTGCGCAC GTTAGCTTTT TGTGGTTTGA GTTGGAGAAC AGTGGGAGGC TCACGGTAGC 360
GTGGAGTGAC GACATTGGGG ATAACACCAG AGGCGTCTTA TCTCCGTTGG ACAAATTATT 420
ATTATGGCTA TGAACATTCA ACATATAATT TAATTAGAAA TTTGCGGATG AAAAAGAGGT 480
AAACAATTGC AGAAATGGTT AAAAATATTA ACGTTGTACA GCAAATGATA ATAAAAAGTG 540
TAACGTACAG TGTGTAAGGA ATGGAAAAAT AATAATTTGG GTTAAAATAA ATATGTAGTT 600
TTCTAACTAT ATAGTACTTT TTGAGAAAAG ATAATATTAT GTGTATTTTT ATTGAAACAA 660
ATAAATGATT TAACAAAAAA AAAAAGAGAA GTTAAAATGA AAAGGAATTA TTATTTTTTA 720
AGTTCTTCCT TCTTTTGTTG GGCCTGTGAC CCTTTTAGTT TTAGTCCACT TCGTTCTCAA 780
AGCTTCAAAA TATTAATTTT GTGACAAACC GACCGGAGCC AACCAAACCG GTTAACATCC 840
TAAAACCAAT CATATTTTAT TAAGTTTTGT GTTGATGCTA AACCAAAAAT CATTGGCATG 900
CATATTTCTA AATTTAGTAA TAAACAAAAA CACTTAGAAA TCACACGTTC ACTATACTAA 960
AAAACGTTGA CAAAAACACA ACAACTATAC TAATAATTAA AGAAGAGAAA ACTGAACCAA 1020
ACTTTTTGTA AACTCCTGAA TTTAAATTAG TAATTGAAGT AAGAAGATGA AGAAGAACAT 1080
GTTAAGCAAA CAAAAAAATT ACACTAAAAT CATATAAAAA TACATAATTA CAAAAGTACC 1140
CATAAGATGG ATTTATTGAT ATGGGTCATC TGTGAAACAA GCCACAGAGA GACAAAGACT 1200
CGTAAGTATT GGGCAACGAA AGCGACCTCC TTTATTCACC ACTGCCATTA ACATGTTCTT 1260
CTTCTCCTTC TTCTTCTACA TTTTATGACC GTTTTACCCT TCAAGAGAGA GAAACAAAAT 1320
CACTCCCTCT CACTCACTCT ATCTCTCTCT TCTGCAAAGC TTCAGAACTC TGGCAGAGAG 1380
ATAAAAGATG ATGGGGTTTT TAACTTTATC CTCCCCAAAT AATTCTTCTT CCCTTCATCT 1440
CTCTCTCTTA CACAACAGGT CCCTACATTT GTACAATCTC CTCTCTTTAA AGACTCTCTC 1500
TCTTTCTCTC TCCATCTCTA TCTTACTCTG TATTTCTGTC GTCTGAGCAC TCAATGAAAC 1560
CACTGTAAAT TTCCGCCAGA ATTTGATGTG ATGGAACGAT AAAAATCATT TTTTCTCGGT 1620
TAAAGTAAAA AAACAAAAAC AAATTTCTGT AGAAATCATA ATAAAAGAAA GAAAAAAAAT 1680
CTAATGTCGG TACATAATAC GGTTCTCTTC TTCTTCTCTA TCCTCTGTTT CTTCTTCATG 1740
GAGACTTGAA AGCTTTTAAA GTATATCTAA AAACGCAGTC GTTTTAAGAC TGTGTGTGAG 1800
AAATGGCTCT GTTTAGAGAT ATTGTTCTTC TTGGGTTTCT CTTCTGCTTG AGCTTAGTAG 1860

CTACTGTGAC TTCAGAGGAG GGTCAGTTAT TATACTGATG CATGCTTCTT CAAGTTCAAG 1920
ATTTTCGTCT TTTTGTTTTA TATTAGTGAA AAAAACTTAA AGATGAGATT TTTATATGAT 1980
TTTTGAAGTT TCATTTGGTG AAAATGAGAT CTGGGTACTT GTTATTTTCT ATTTTTGCTT 2040
TTTGTAATGG TTTTTTTTTA CTTGGTGGGT CTTCTATAGA ATCAAAAGAA GCTTTGAATA 2100
AATTAGGGTT TGAGTTTTAT TTTGTTTTCT TGGAAGTTGA ATTTTTAATC TTCTCAAGAA 2160
CTGACAAATA TTTTTTTTTG TTTTTGTGCG TGTGTGTTAA TAAAATATCC TTAAAACAAA 2220
ATTAAAGGAG CAACGTTGCT GGAGATTAAG AAGTCATTCA AAGATGTGAA CAATGTTCTT 2280
TATGACTGGA CAACTTCACC TTCTTCGGAT TATTGTGTCT GGAGAGGTGT GTCTTGTGAA 2340
AATGTCACCT TCAATGTTGT TGCTCTGTAA GTTTCTTCAT TCCTTTAGAT TACTATTACA 2400
GTGGTTTTTG GTGTTCTTGT GGGAAAAAGT TGTAATTTGT TTTGTGTGTG TTTTCTATGT 2460
TTTGTAGTAA TTTGTCAGAT TTGAATCTTG ATGGAGAAAT CTCACCTGCT ATTGGAGATC 2520
TCAAGAGTCT CTTGTCAATG TAACTGTTTC AACATTCACT GTAGCATGAA ATAAAGTATC 2580
TTACTTTAAT TCTATTCCAC TCTCTGAGTT GTGACTTTTG TCTTCTGTTT TTTTCTAATG 2640
TAGTGATCTG CGAGGTAATC GCTTGTCTGG ACAAATCCCT GATGAGATTG GTGACTGTTC 2700
TTCTTTGCAA AACTTGTAAG AACAGTGATT GGTGTTATTC TACCATTAAA CTTTTGTTCA 2760
TAGAGGTTTT ATTTGATGAA GTGTGTTCAT GTTGTTTTTA ATTCAGAGAC TTATCCTTCA 2820
ATGAATTAAG TGGTGACATA CCGTTTTCGA TTTCGAAGTT GAAGCAACTT GAGCAGCTGT 2880
AAGTAGCTAG TTATTCTGCT ACTAGTCTTC ATATGTCATT GCTAAAAATA TACTCACCAT 2940
GTGGAATATG GATTTTTACT TTGTCCAGGA TTCTGAAGAA TAACCAATTG ATAGGACCGA 3000
TCCCTTCAAC ACTTTCACAG ATTCCAAACC TGAAAATTCT GTATGTTCCC CATGATTCTT 3060
ACATGTCTTA CTACTTTTAG CTATATAGGT GATCATACAT GTGTAATTTC AATTGCAGGG 3120
ACTTGGCACA GAATAAACTC AGTGGTGAGA TACCAAGACT TATTTACTGG AATGAAGTTC 3180
TTCAGTATCT GTAAGTGTCA ATGTTTTTTG AAGTCTGTCA ATGTCTCTTC ATTACCCGGT 3240
GATAATTGTT GTACTATGAT GAGCAGTGGG TTGCGAGGAA ACAACTTAGT CGGTAACATT 3300
TCTCCAGATT TGTGTCAACT GACTGGTCTT TGGTATTTGT GAGTCTTCTT GCACATCTGA 3360
ATAGTATGAT GAGTTCTTTT GTAAATATCA AATATCTGAC TTTGTTTTGA TATTGAATCA 3420
GTGACGTAAG AAACAACAGT TTGACTGGTA GTATACCTGA GACGATAGGA AATTGCACTG 3480
CCTTCCAGGT TTTGTATGTG CCTCTTTCTC TACTTCTAAA CATCATTACT GTAATTTGGG 3540
TTACTTAAGA AAATCTACTT AACTGGTTTG CTTATTACGA ACTCAGGGAC TTGTCCTACA 3600
ATCAGCTAAC TGGTGAGATC CCTTTTGACA TCGGCTTCCT GCAAGTTGCA ACATTGTTAG 3660
TTCTCACCTC TACTAATCTT TTGCTTTAAA TTTTGGCTAG CCTTTGTTTT CTTTTAAAGA 3720
AGATCATTTT CTTATCTTAG ATCATTGCAA GGCAATCAAC TCTCTGGGAA GATTCCATCA 3780
GTGATTGGTC TCATGCAAGC CCTTGCAGTC TTGTAAGTAC TTTTCTTCTA ATCAATGAAG 3840
CTACTTATAA CATTTTCATG AACTTAGGTT ATATGTTTTC TTTTACAGAG ATCTAAGTGG 3900
CAACTTGTTG AGTGGATCTA TTCCTCCGAT TCTCGGAAAT CTTACTTTCA CCGAGAAATT 3960
GTAATTCTTT ACCTGTTTGT TTTCAGTTTG GAGTCAAATG TCATACCATG TTAATGATAG 4020
TGATTTATCT TTTTGGCTTT ATCTCTAGGT ATTTGCACAG TAACAAGCTG ACTGGTTCAA 4080
TTCCACCTGA GCTTGGAAAC ATGTCAAAAC TCCATTACCT GTATGACCAA CCTTCTCTTC 4140
ACTTCTCTTT TTGCATACAG TCACTACTAA GTTGTGTTTC CTTATCAACT ATTTGTAAAA 4200
TATTCATAGG GAACTCAATG ATAATCATCT CACGGGTCAT ATACCACCAG AGCTTGGGAA 4260
GCTTACTGAC TTGTTTGATC TGTAAGTAGT TCTTCCTATG CTTGACATGT TTTGATGTTC 4320
TTATGCTTAT ATGAACTATG TACATATAGG AATGTGGCCA ACAATGATCT GGAAGGACCT 4380
ATACCTGATC ATCTGAGCTC TTGCACAAAT CTAAACAGCT TGTATGTATC TCTTTCTCTG 4440
AAAACTTCTC ACTTGAATGT TCAAGATTGG TGCTTTATAT GATTTTGTGT CTCATTAATG 4500
TAATGTAGAA ATGTTCATGG GAACAAGTTT AGTGGCACTA TACCCCGAGC ATTTCAAAAG 4560
CTAGAAAGTA TGACTTACCT GTAAGTATCG ACGCTGAGAA TTTCTCTAAT CTTATATAAT 4620
ATATAGTTCC ACAGCGTTTG TTTTTTCGAA TTTCAAGTCA TTAACTACTG AGTTTTTGGT 4680
TGCCTTTGAT TTATCGGTTC AACCAGTAAT CTGTCCAGCA ACAATATCAA AGGTCCAATC 4740
CCGGTTGAGC TATCTCGTAT CGGTAACTTA GATACATTGT AAGTGTTTCT TGTTTTCTGT 4800
GAAGTATACA TCATTATATG TGCCTTGTCT CACATTTATT AAATTTAATG ACATTTGAAG 4860

GGATCTTTCC AACAACAAGA TAAATGGAAT CATTCCTTCT TCCCTTGGTG ATTTGGAGCA 4920
TCTTCTCAAG ATGTGAGCAT CCATAAGACC TCCAGTTTTA TTGTTTATTT CTAGCAAAAG 4980
ATGAAAATGG TTTGTGAACT CTTGCATTCT TGTTATAGGA ACTTGAGTAG AAATCATATA 5040
ACTGGTGTAG TTCCAGGCGA CTTTGGAAAT CTAAGAAGCA TCATGGAAAT GTAAGAAGTT 5100
AACTTCTATC TGCTTGGTTA GAGTTTTTTT CATTTATCTC AATTACTGTT CTGAATTTGT 5160
GTGTTTGTGG TTGCAGAGAT CTTTCAAATA ATGATATCTC TGGCCCAATT CCAGAAGAGC 5220
TTAACCAATT ACAGAACATA ATTTTGCTGT AAGCAATCTT CCTCTTATCC CTTCCAAGCT 5280
GTTAAGAAAT TGTTTTTGTA GAATGAAACT AAAACTCTGT ATACACAATA ATGAGGTCAC 5340
TATAGTGTGA TCCAGGAACA TGTATTGGGT TGGTGATCTA TCTAATGTTG TGTTTCTTAA 5400
AATTGCTTGC AGGAGACTGG AAAATAATAA CCTGACTGGT AATGTTGGTT CATTAGCCAA 5460
CTGTCTCAGT CTCACTGTAT TGTAAGTAGG CACCTTTGGT TCTGAAACAT TTTTTGTCCC 5520
TCTTTGTGCA TCTTTTGCTA AGAATATAAC CCTGCAATCT TCACTAACTC TTATAGGAAT 5580
GTATCTCATA ACAACCTCGT AGGTGATATC CCTAAGAACA ATAACTTCTC AAGATTTTCA 5640
CCAGACAGGT ATGGTAATTT AGCAGGTTTT GGTATTGTGC ATTTTGTTTT GTTTGCTAAT 5700
ATCTATGTTT ATGTTTTTGG ATAAAGCTTC ATTGGCAATC CTGGTCTTTG CGGTAGTTGG 5760
CTAAACTCAC CGTGTCATGA TTCTCGTCGA ACTGTACGAG GTGATTACAT TCTTCTAAAA 5820
GCTTCCATTC ACAAAACCTA AGATAATTAA AGCTCATGTT TCTATCCATG TTTTGTCTGC 5880
AGTGTCAATC TCTAGAGCAG CTATTCTTGG AATAGCTATT GGGGGACTTG TGATCCTTCT 5940
CATGGTCTTA ATAGCAGCTT GCCGACCGCA TAATCCTCCT CCTTTTCTTG ATGGATCACT 6000
TGACAAACCA GGTCTACTCT CCAAACCACT TTACGAATGT TCTTCACCTA CAATGTAATC 6060
CAATAGTTAA TCCTTAAATT TCCTGGTGAC ATCAGTAACT TATTCGACAC CGAAGCTCGT 6120
CATCCTTCAT ATGAACATGG CACTCCACGT TTACGAGGAT ATCATGAGAA TGACAGAGAA 6180
TCTAAGTGAG AAGTATATCA TTGGGCACGG AGCATCAAGC ACTGTATACA AATGTGTTTT 6240
GAAGAATTGT AAACCGGTTG CGATTAAGCG GCTTTACTCT CACAACCCAC AGTCAATGAA 6300
ACAGTTTGAA ACAGAACTCG AGATGCTAAG TAGCATCAAG CACAGAAATC TTGTGAGCCT 6360
ACAAGCTTAT TCCCTCTCTC ACTTGGGGAG TCTTCTGTTC TATGACTATT TGGAAAATGG 6420
TAGCCTCTGG GATCTTCTTC ATGGTAAGTC TCATCGCCAA ACATAGAAAA TTATTTGAAT 6480
CTTCTGTGAC ATAACAACTT GCTTGTGTGT TTTGTAAAGG CCCTACGAAG AAAAAGACTC 6540
TTGATTGGGA CACACGGCTT AAGATAGCAT ATGGTGCAGC ACAAGGTTTA GCTTATCTAC 6600
ACCATGACTG TAGTCCAAGG ATCATTCACA GAGACGTGAA GTCGTCCAAC ATTCTCTTGG 6660
ACAAAGACTT AGAGGCTCGT TTGACAGATT TTGGAATAGC GAAAAGCTTG TGTGTGTCAA 6720
AGTCACATAC TTCAACTTAC GTGATGGGCA CGATAGGTTA CATAGACCCC GAGTATGCTC 6780
GCACTTCACG GCTCACTGAG AAATCCGATG TCTACAGTTA TGGAATAGTC CTTCTTGAGT 6840
TGTTAACCCG AAGGAAAGCC GTTGATGACG AATCCAATCT CCACCATCTG GTTTGTTCTT 6900
TCTTGCCTAT CTCTCTCAGC TGCTCTGTTT AGGTCAAGTC CGTAATCTTG TTTTCATTGA 6960
TTCACTTACA TCAGATAATG TCAAAGACGG GGAACAATGA AGTGATGGAA ATGGCAGATC 7020
CAGACATCAC ATCGACGTGT AAAGATCTCG GTGTGGTGAA GAAAGTTTTC CAACTGGCAC 7080
TCCTATGCAC CAAAAGACAG CCGAATGATC GACCCACAAT GCACCAGGTG ACTCGTGTTC 7140
TCGGCAGTTT TATGCTATCG GAACAACCAC CTGCTGCGAC TGACACGTCA GCGACGCTGG 7200
CTGGTTCGTG CTACGTCGAT GAGTATGCAA ATCTCAAGAC TCCTCATTCT GTCAATTGCT 7260
CTTCCATGAG TGCTTCTGAT GCTCAACTGT TTCTTCGGTT TGGACAAGTT ATTTCTCAGA 7320
ACAGTGAGTA GTTTTTCGTT AGGAGGAGAA TCTTTAAAAC GGTATCTTTT CGTTGCGTTA 7380
AGCTGTTAGA AAAATTAATG TCTCATGTAA AGTATTATGC ACTGCCTTAT TATTATTAGA 7440
CAAGTGTGTG GTGTGAATAT GTCTTCAGAC TGGCACTTAG ACTTCCTATA AGTTCTTGCC 7500
TATCTAAGTT TTTCTAAATT GGGTTATTCT TGTAACATAT CTTAGATCTA GTACTCAACA 7560
CCACGTCACC ACCACAAAAG ATTTCTTATG CTCAAAAACA TATACATAGA AAGAACCTTC 7620
TAAACTACGA GAAACGTTTT GCTATGTAGT GTTATATGTC AACCACGTCT ATGAGAGTGC 7680
AAACGATAGG TTAATAAGTT TTCTCACTTG GCAATAAAAA TGATAAACAA ATATATTGTC 7740
TGATTAATTT ATTTTATATA GTTTTTTTAT AATTTCTTAT ATTAATTCGA ACTCATACAG 7800
CGCGTGAGAC TTTCTAGTTT AGTATAAAGT ACGTATTTTT GCAAAATCAA AATCGTAAAT 7860

ACATACATTT TAAAATGTTA AAAAAGATAA ATCCGTACAC CATTTAAAAA TGGCATTTTC 7920
CTAAGATTTT TTTCAAAAAA GGCATTTTAG ACAAGAACTA ATTACTACAA CTAAAATCTA 7980
CTAACTTTGG TTTTTATGTA TACATTTACG AGAGTCTACA CAAAAAAAAT ACATAAAAGA 8040
AGAAGTAGTA AATAATTAAA ACGTAAAAAA AAAGACTTTT CAAGAAGGCA GAAGAGTAGC 8100
ACTGTTGTGC GATTGTAAAA TCGTCTTGAT TGTTGTTTAT CCCACTGATA AGCCTACCCT 8160
TTTCAAAACT TGTTCTAAGT TTAAATTCTA TTTTTGAACA TGACATACAG TATAAGGCTT 8220
TTTAAAGATA TCATCTTGAT TTTGTTTCTT CCACAGGGAA GCCCTATCCT TTCTTACATA 8280
ATCTTTGTTA GATAATTTTT TATTATTTTC AAAAAAAATA AAATTGAACA TAAGTTTTCT 8340
CAAAGTAATA TGTTCTAACA ATAATAAACA TAATATCATT TTTTTGTTTT AAACTATAAA 8400
GGACTAACAT GGTAAAAAGT TGCAATATAT AAATGATAAT TTAAACTAAA AATTAGAATA 8460
TGGTAACTTT TTCTTCAACA ACATGCCACA TTCGGCTACA TGTCCACTAG GAAGTGTTAT 8520
TATAGAATCG TTAATGTTGG GTACGCTTAT GAAATTATCA ATGTTTGCTT AAATCTATGC 8580
TTAGAAAATT ACCAATATTA CCTTAAAACT ATATTTACGA ATGACCAATA TTGCTTAGAA 8640
CTATGCTTAT GAAATTACCA ATATTTTCTT AAAACTTAAA CACAAAACTC TTTAACAAAA 8700
AAAACTTTAT TTTTATTTTT ATTTTTTTGG CAAAAAAAAA AACTTTATTT ATAAAGTGAA 8760
AGTCTCCAGA TAATTTTGAA TTTCATTTTT CCAGTTTTTA TTTAGAATAA TTTTTCTTCA 8820
TTTACAAAAT AAAAGAAAAC CCTAGGGTTT AGGGTTTAGG GTTTAGGAAA AAGCGATGAT 8880
ATATTAATTG TTATGAAATG TTTTTTTAAA AATAGTTAAC CAAACATTTT TTTAAAGAGA 8940
GTTTAGTTTC ACAAGGCATT TGTAAATTAG AGTAATTATC AATAAAAATG GAAGACAATC 9000
TAATTATTAT TTAGCAAAAA CTATATTTAG GAAAATTAGT TAAAGTTTAG AAATATATCA 9060
TCATAGTGTC AAACTAATTA AAATTATTTA ATTTTGTGAT ATACGTGATC ATATAATTTT 9120
ATGAATATTT AATATTATGA TACATGTAAC TCAGTAAACC TAAATTTAGA AGAAAAGTCA 9180
AAATAATCAT AACCAATTTA GATTCAACTT CTACTTTTGT TCCAAGAAAA AAACACATGG 9240
TTTGTTTTGT GGGATACTAA TGACATCTAT CAAAATCTAT GAAACCAAAT CTAGA 9295

配列番号：3
配列の長さ：18
配列の型：核酸
鎖の数：一本鎖
トポロジー：直鎖状
配列の種類：他の核酸　合成DNA
アンチセンス：NO

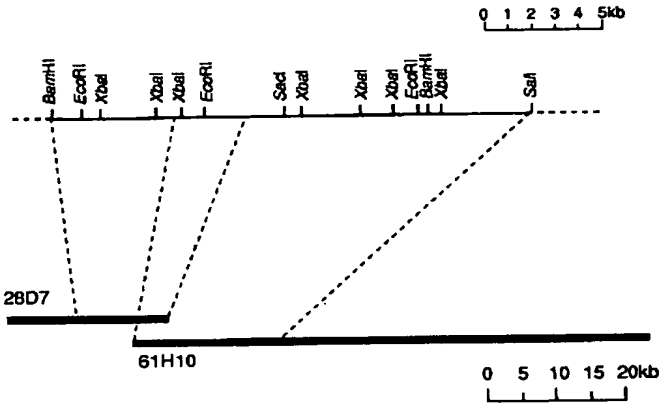
配列：
TATCTAAAAA CGCAGTCG 18

配列番号：4
配列の長さ：18
配列の型：核酸
鎖の数：一本鎖
トポロジー：直鎖状
配列の種類：他の核酸　合成DNA
アンチセンス：YES

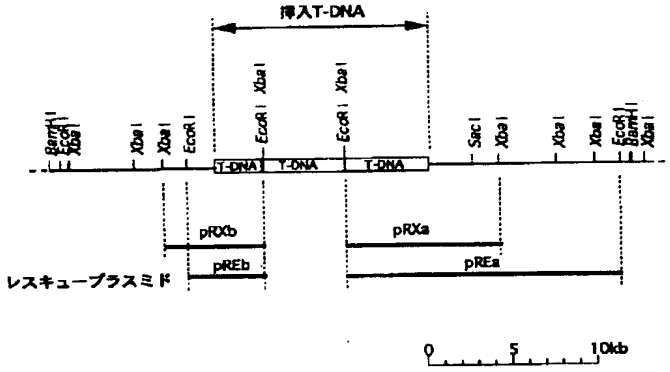
配列：
AAGATTCTCC TCCTAACG 18

【図面の簡単な説明】
【図１】　実施例で得られた形態制御遺伝子クローンの制限酵素地図。
【図２】　制限酵素地図及びサザン解析から推定された染色体ＤＮＡ断片上のT-DNA挿入部位を示す図。
【図３】　T-DNA挿入部位近傍配列のサブクローンの位置を示す図。

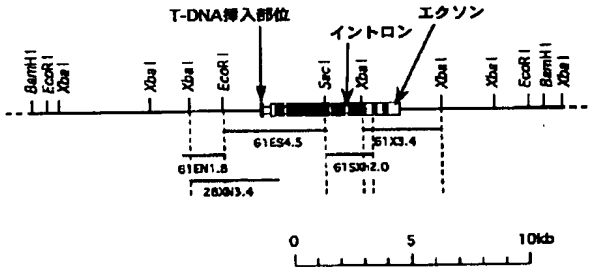
【図1】


0 1 2 3 4 5kb
BamHI
EcoRI
XbaI
XbaI
XbaI
EcoRI
SacI
XbaI
XbaI
XbaI
EcoRI
BamHI
XbaI
SalI
28D7
61H10
0 5 10 15 20kb

【図2】


挿入T-DNA
T-DNA
T-DNA
T-DNA
pRXb
pREb
pRXa
pREa
レスキュープラスミド
0 5 10kb

【図3】


T-DNA挿入部位
イントロン
エクソン
61ES4.5
61EN1.8
61SX3.4
61SN2.0
28XN3.4
0 5 10kb

フロントページの続き

(72)発明者　ロバート　エフ．　ウィッティア
茨城県つくば市千現２－１－６　つくば研究支援センターＤ－６　株式会社三井業際植物バイオ研究所内

EXHIBIT E

The Plant Cell, Vol. 8, 735–746, April 1996 © 1996 American Society of Plant Physiologists

The Arabidopsis *ERECTA* Gene Encodes a Putative Receptor Protein Kinase with Extracellular Leucine-Rich Repeats

Keiko U. Torii,[a,1,2] Norihiro Mitsukawa,[b,2] Teruko Oosumi,[b,3] Yutaka Matsuura,[c] Ryusuke Yokoyama,[c] Robert F. Whittier,[b] and Yoshibumi Komeda [a,c]

[a] Molecular Genetics Research Laboratory, University of Tokyo, Hongo, Tokyo 113, Japan
[b] Mitsui Plant Biotechnology Research Institute, TCI-D21, Sengen, Tsukuba 305, Japan
[c] Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060, Japan

Arabidopsis Landsberg *erecta* is one of the most popular ecotypes and is used widely for both molecular and genetic studies. It harbors the *erecta* (*er*) mutation, which confers a compact inflorescence, blunt fruits, and short petioles. We have identified five *er* mutant alleles from ecotypes Columbia and Wassilewskija. Phenotypic characterization of the mutant alleles suggests a role for the *ER* gene in regulating the shape of organs originating from the shoot apical meristem. We cloned the *ER* gene, and here, we report that it encodes a putative receptor protein kinase. The deduced ER protein contains a cytoplasmic protein kinase catalytic domain, a transmembrane region, and an extracellular domain consisting of leucine-rich repeats, which are thought to interact with other macromolecules. Our results suggest that cell–cell communication mediated by a receptor kinase has an important role in plant morphogenesis.

INTRODUCTION

The form of higher plants is the consequence of the repetitive divisions and subsequent differentiation of the cells produced by the shoot apical meristem. The shoot apical meristem keeps initiating new organs throughout the life of plants, while maintaining itself as a formative region. Organ primordia are derived from numerous cells that originate from multiple lineages (Szymkowiak and Sussex, 1992). These cells coordinate their growth patterns to develop determinate organs. Thus, cell–cell signaling is crucial in determining organ shape.

The molecular nature of these signals for cell–cell communication is not fully understood. Recent molecular genetic studies using Antirrhinum and maize have, however, identified genes that potentially mediate cell–cell communication. Mosaic analyses using the maize leaf mutants *Teopod*, *Rough sheath*, and *Knotted* indicate that the gene products may act non–cell autonomously (Sinha and Hake, 1990; Dudley and Poethig, 1993; Becraft and Freeling, 1994). KNOTTED may be able to move from mesophyll (L2) cells, which express *KNOTTED*, to the epidermal (L1) cells, which do not express *KNOTTED* (Jackson et al., 1994). In Antirrhinum, a mosaic analysis of *floricaula* has also shown that it can act non–cell autonomously in the floral meristem (Carpenter and Coen, 1995; Hantke et al., 1995). It is not known how *Teopod*, *Rough sheath*, or *floricaula* signal to the surrounding cells to coordinate organogenesis.

We have taken a genetic approach to determine the mechanism specifying organ shape. At the vegetative stage, an Arabidopsis shoot apical meristem produces leaves and axillary meristems. Upon entering the reproductive stage, the shoot apical meristem converts into the inflorescence meristem, which then produces floral meristems. In the typical rosette plant Arabidopsis, a transition from vegetative to reproductive development accompanies the elongation of the inflorescence stalk (i.e., bolting). Generation of floral buds and subsequent elongation of the internodes seem tightly coupled, thus producing a highly ordered branching pattern. We have performed a mutant screen for altered inflorescence branching patterns (Tsukaya et al., 1993; Komeda and Torii, 1994) and isolated five mutations allelic to Landsberg *erecta* (*Ler*).

Ler is one of the most popular ecotypes of Arabidopsis and has been widely used for both molecular and genetic studies (Hwang et al., 1991; Anderson and Mulligan, 1992). *Ler* was isolated from mutagenized seed populations in the 1950s (Rédei, 1992). It harbors the *erecta* (*er*) mutation and shows an altered organ shape. *Ler* develops a very compact inflorescence with flowers clustering at the top. *Ler* plants also display round leaves with short petioles and short and blunt siliques (Rédei, 1992; Bowman, 1993). The compact stature of *Ler* is preferred by genetists, and thus *Ler* has been used as a wild-type strain to isolate numerous mutants. These

[1] To whom correspondence should be addressed. Current address: Osborn Memorial Laboratories, Department of Biology, Yale University, 165 Prospect Street, New Haven, CT 06520-8104.
[2] These authors contributed equally to the work.
[3] Current address: U.S. Department of Agriculture, Agricultural Research Service, 800 Buchanan Street, Albany, CA 94710.

include mutants in photomorphogenesis, phytohormone biosynthesis and signal transduction, and flower organ identity. Many such mutants have been characterized, despite the fact that no one knew the nature of the *er* mutation.

As an initial step toward understanding the molecular mechanism regulating the specific organ shape, we isolated the *ER* gene. The *ER* gene encodes a putative receptor protein kinase with an extracellular ligand binding domain, implicating the existence of an intercellular signal transduction pathway that is required for proper development of organs derived from the shoot meristem.

RESULTS

Phenotypes of *er* Mutants

We isolated new mutant alleles at an *er* locus from ecotypes Columbia (Col) and Wassilewskija (WS). Plants homozygous for all *er* alleles show significantly compact inflorescences compared with those of the wild types (Figure 1A). Inflorescence stems are thicker in plants homozygous for *er* alleles when compared with the wild types (data not shown). It seems that the short and thick inflorescence stem phenotype makes *er* mutants "erect." Flower buds are clustered at the top of the inflorescence in plants homozygous for each *er* allele without affecting phyllotaxis (Figure 2). Moreover, the number of flower buds at the first flowering was increased in *er* mutants (see the legend to Figure 2). Thus, the *er* mutation may somehow affect the coordination of stem elongation and flower bud formation. The number of lateral inflorescences was not altered in plants homozygous for all *er* alleles, suggesting that the *er* mutation may not affect apical dominance (data not shown). Siliques are blunt, short, and wider in plants homozygous for all *er* alleles (Figures 1C, 1D, and 3). Because flowers of *er* and wild-type plants are similar (Figure 2), it is likely that the *er* mutation affects the elongation of carpels after fertilization. *er* mutants also have very short pedicels compared with those of the wild types (Figure 1B).

Leaf morphology is varied in plants homozygous for each *er* allele (data not shown). *Ler* has round leaves with a short petiole, as previously described (Rédei, 1992; Bowman, 1993). Leaves of *er-102* are small and curly, and these traits cosegregate with the other phenotypes described above (data not shown). In contrast, leaves of *er-101* and *er-104* seem less affected by the mutations, and leaves of *er-103* are almost indistinguishable from those of the wild types (data not shown). The leaves of *er-101/er-102* heterozygous plants display an intermediate phenotype (data not shown). No difference was observed in roots of wild-type plants and *er* mutants (data not shown). From these observations, a corresponding wild-type *ER* gene is likely to be required for proper elongation of various organs of shoot meristem origin.

Precise phenotypic analysis revealed the degree of severity among *er* alleles (Figures 1A to 1D). *er-102* plants are the most compact (Figure 1A), having the shortest petioles (Figure 1B) and the shortest and widest siliques (Figures 1C and 1D). Thus, we conclude that this is the most severe allele. In contrast, *er-103* plants are taller (Figure 1A), having long siliques (Figure 1C) of the same width as those of the wild types (Figure 1D). Therefore, *er-103* seems to be the weak allele. The severity of the *er-101*, *er-103*, *er-105*, and *Ler* alleles seems to be approximately the same (Figures 1A to 1D), except that *Ler* plants tend to have longer pedicels (Figure 1B).

Molecular Identification of the *ER* Locus and Defects in the *er* Alleles

The *er-104* allele, generated by T-DNA insertional mutagenesis, was used to isolate the *ER* gene. Initially, *er-104* harbored two independent T-DNA insertions. We backcrossed *er-104* twice into wild-type Col and obtained a line that has a single T-DNA insertion. Genetic analyses indicated that the insertion is tightly linked to the *er* locus (data not shown). By using the T-DNA as a probe, DNA gel blot analysis revealed a complex insertion of 2.5 T-DNA copies with left borders at both genomic DNA junctions (data not shown). By using the pBR322-derived replication origin and ampicillin-resistance marker present in this portion of the T-DNA, both genomic junctions were recovered separately by plasmid rescue (Figure 4A). They showed a polymorphism between *er-104* and the WS wild type (Figure 4B) and hybridized with six of eight yeast artificial chromosome clones (namely, EG1D5, EG2A1, EG2B1, EG10A10, EG10H3, and EG16C6), which contain the *GPA1* locus, an Arabidopsis G protein α subunit gene located within 1 centimorgan of the *er* locus (data not shown) (Ma et al., 1990; Hwang et al., 1991; Hwang and Goodman, 1995).

Two independent transcripts of 3.3 and 0.6 kb were found within the 13-kb region used to screen the cDNA library (Figure 4A). RNA gel blot analysis revealed the absence of the 3.3-kb transcript in *er-104* and *er-105*, whereas the expression level of the 0.6-kb transcript was not affected in these alleles (data not shown), suggesting that the 3.3-kb transcript is most likely the *ER* transcript.

Six cDNA clones corresponding to the 3.3-kb transcript were isolated from an Arabidopsis Col cDNA library. The longest cDNA is 3176 bp and contains a single open reading frame of 976 amino acid residues with a calculated molecular mass of 107.3 kD (Figure 5). The presence of an in-frame stop codon upstream of the first ATG confirmed that this is the initiation codon. Comparison of genomic and cDNA sequences revealed the presence of 26 introns (Figures 5 and 6A). Comparison of cDNA and plasmid-rescued DNA revealed that the T-DNA was inserted in the 5′ untranslated region (Figure 5). The T-DNA insertion was associated with the deletion of 28 nucleotides, from −6 to +22 of the 5′ terminus of the full-length *ER* cDNA (Figure 5), suggesting that the transcriptional initiation point was deleted. *er-105* was generated by fast-neutron irradiation

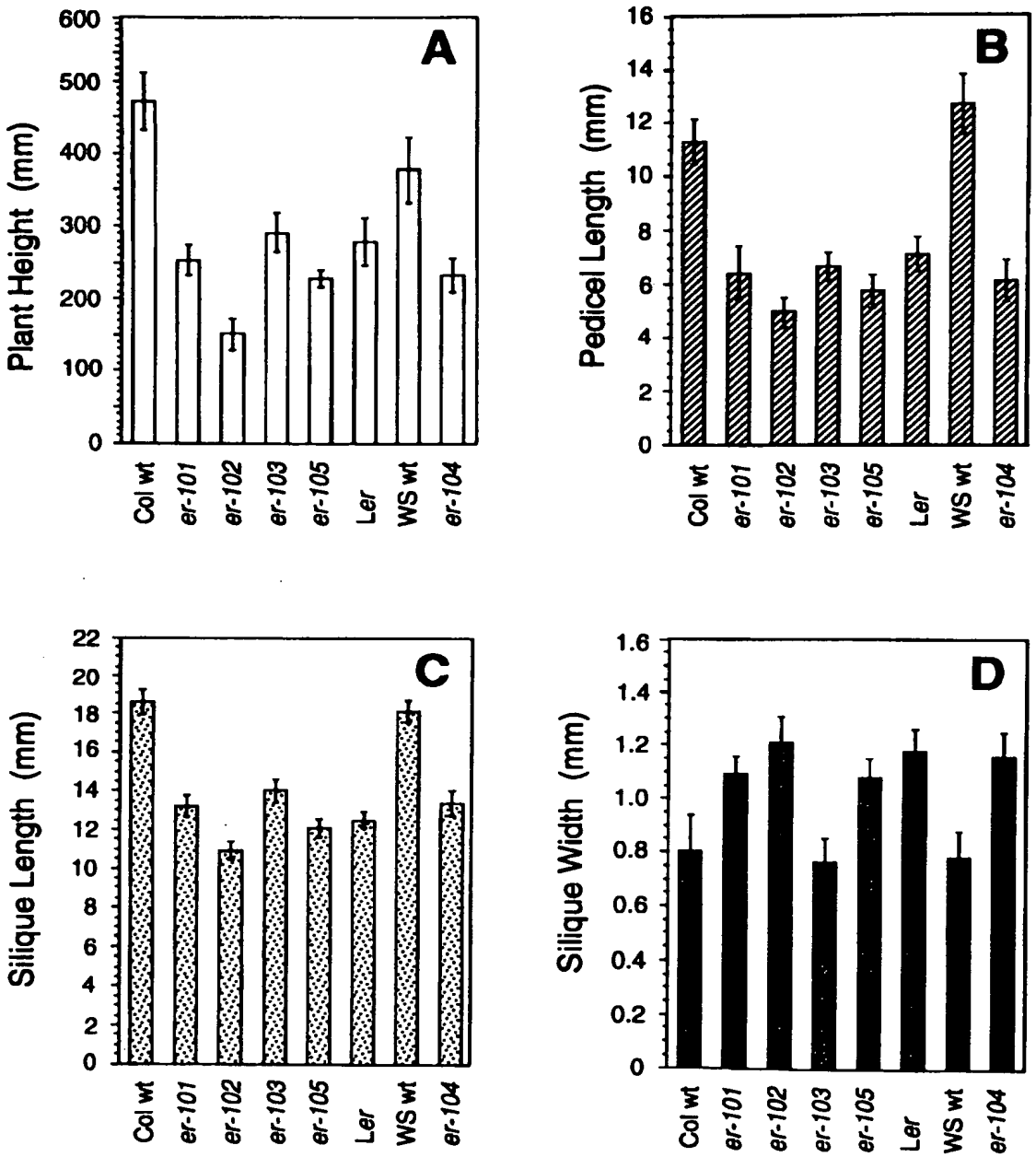


Figure 1. Comparison of the Inflorescence, Pedicel, and Silique Lengths and the Silique Widths of *er* Mutants and Wild-Type Plants.

(A) The length of the main inflorescence of 40-day-old wild-type (wt) plants and plants homozygous for each *er* allele. Length was measured and designated as plant height. At least 27 plants with each *er* allele and the wild-type plants were measured.

(B) The pedicel length of 40-day-old plants. Ten pedicels from the base of the main inflorescence of five individual plants (total of 50 pedicels) with each *er* mutant allele and the wild-type plants were measured.

(C) Silique length. Ten fully expanded siliques from five individual plants (total of 50 siliques) with each *er* allele and the wild-type plants were measured before desiccation.

(D) Silique width. Fifty siliques, as given above for **(B)** and **(C)**, were measured.

The mean values are shown. Error bars represent standard deviation.

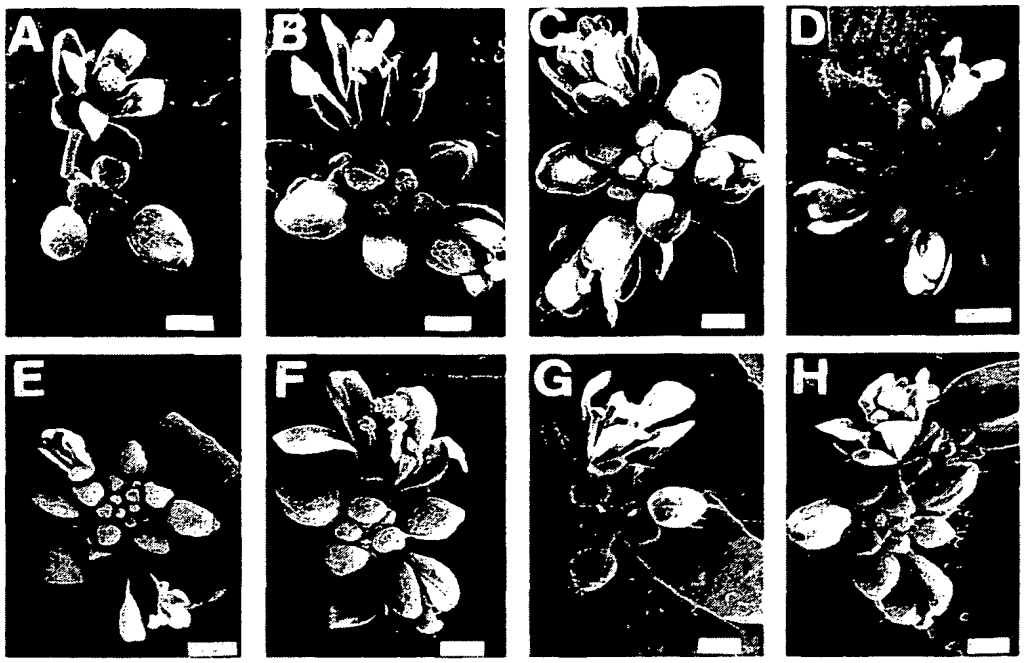


Figure 2. Inflorescence Morphology of *er* Mutants and Wild-Type Plants.

Shown are top views of inflorescences at first flowering.
(A) Wild-type Col.
(B) *er-101*.
(C) *er-102*.
(D) *er-103*.
(E) *er-105*.
(F) *Ler*.
(G) Wild-type WS.
(H) *er-104*.
Flower buds of plants homozygous for all *er* alleles are tightly clustered at the top compared with those of the wild-type plants. Numbers of flower buds at first flowering were 19, 21, and 21 for *er-101*, *er-102*, and *er-104*, respectively, and 15 and 16 in Col and WS wild types, as shown. Bars = 500 μm.

and was therefore expected to result from a gross DNA rearrangement. We performed DNA gel blot analysis with *er-105* (Figure 4C) and found that ~4 kb of DNA of unknown origin is inserted within the *ER* locus (Figures 4A and 4C). A precise polymerase chain reaction analysis determined the region of insertion between +5 and +1056 after the first ATG for translation in the genomic sequence (data not shown). Molecular defects in the *er-104* and *er-105* alleles are consistent with the absence of the transcripts (Figure 7A). None of the other *er* alleles or *Ler* showed polymorphism with wild-type Col or WS when their genomic DNA was analyzed by DNA gel blotting (data not shown). The *ER* gene is most likely a single copy (Figure 4C).

To confirm further that we had cloned the *ER* gene, two additional alleles were characterized at the molecular level. Ethylmethane sulfonate–generated *er-103* has a G-to-A transversion at position +846, which changes amino acid 282 from methionine to isoleucine (Figures 6A and 6B). In *Ler*, a T-to-A transversion at position +2249 was found, and this change results in a substitution of lysine for isoleucine at amino acid 750 (Figures 6A and 6C). *Ler* also contains two silent mutations (T-to-C transversion at positions +1389 and +1608), which are most likely due to the polymorphism between Col and *Ler* ecotypes.

The *ER* Gene Encodes a Putative Receptor Protein Kinase with a Ligand Binding Domain

The deduced amino acid sequence of *ER* shows characteristics of a transmembrane receptor protein kinase with distinct domains (Figures 6A to 6C). Two hydrophobic domains are

present at the N terminus (amino acids 1 to 20) and between amino acids 580 and 602 (Figures 5 and 6A). These are consistent with a signal peptide and a transmembrane domain, respectively (Weinstein et al., 1982; von Heijne, 1983). The C-terminal cytoplasmic region (amino acids 648 to 914) comprises a putative catalytic domain of protein kinase (Figures 5 and 6C) (Hanks and Quinn, 1991). A putative extracellular domain (amino acids 75 to 530) contains 20 tandem copies of a 24–amino acid leucine-rich repeat (LRR) (Figures 5 and 6B). These repeats have been implicated to play a role in protein–protein interactions (Kobe and Deisenhofer, 1994). Each unit of the LRR is encoded by identically sized exons, and introns of similar sizes are present at the exact same position, between the second and third nucleotides of the codon for leucine at position 13 (underlined) in the consensus P**LG-*L**L**L*L**N*L*G*I (asterisks represent nonconserved amino acids) (Figures 5 and 6B). Thus, this domain has most likely evolved by exon duplication. The mutation in *er-103* occurs in a consensus position of the LRR, changing methionine to isoleucine (Figures 6A and 6B). Both amino acids are similar, but the latter lacks sulfur. Thus, the mutation may alter the structure of the LRR domain slightly, possibly affecting the receptor–ligand interaction and resulting in the weak phenotype. The presence of 15 N-glycosylation sites (Asn-X-Ser/Thr) (Figure 5) suggests that ER is a glycosylated protein.

The protein kinase domain of *ER* has all 11 conserved subdomains of eukaryotic protein kinases and all invariant amino acid residues in their proper positions (Hanks and Quinn, 1991). This domain of *ER* is most closely related to the predicted receptor-like protein kinases (RLKs) in higher plants: maize ZmPK1 (36% identity; Walker and Zhang, 1990), Brassica SRK6 (32% identity; Stein et al., 1991), and Arabidopsis TMK1 (35% identity; Chang et al., 1992) and RLK5 (40% identity; Walker, 1993). ER appears to fall into the serine/threonine class of protein kinases, because it contains diagnostic sequences of this family (subdomain VIb and VIII) (Hanks and Quinn, 1991), and SRK6, TMK1, and RLK5 are demonstrated to have serine/threonine substrate specificity (Chang et al., 1992; Stein and Nasrallah, 1993; Horn and Walker, 1994). A point mutation in L*er* changes isoleucine, a highly conserved amino acid

Figure 3. Morphology of Fully Expanded Siliques of *er* Mutants and Wild-Type Plants.

(A) Wild-type Col.
(B) *er-101.*
(C) *er-102.*
(D) L*er.*
(E) Wild-type WS.
(F) *er-104.*
Siliques from plants homozygous for all *er* alleles are blunt and short compared with those of the wild-type plants. Bar in **(A)** = 1 mm.

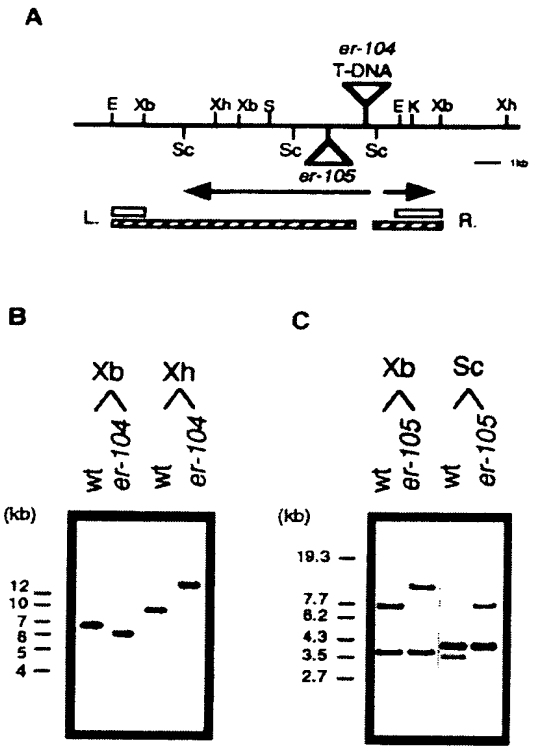


Figure 4. Structure of the *ER* Region and DNA Gel Blot Analysis.

(A) Shown are the restriction map and genomic structure of the *ER* locus and its flanking region. Fragments used to screen yeast artificial chromosomes and genomic P1 libraries are indicated by the open bars, and fragments used to screen the cDNA library are indicated by the hatched bars. Arrows indicate the orientation and lengths of transcripts from 5′ to 3′. The insertion site of T-DNA in *er-104* and the insertion site of DNA of unknown origin in *er-105* are also indicated. E, EcoRI; K, KpnI; S, SacI; Sc, ScaI; Xb, XbaI; Xh, XhoI.

(B) DNA gel blot analysis using plasmid-rescued fragments as a probe revealed a polymorphism between wild-type WS (wt) and a T-DNA–tagged allele, *er-104*. Probe R was used for hybridization. One microgram of total genomic DNA was digested with XbaI (Xb) and XhoI (Xh). T-DNA, originating from the transformation vector pGDW32 (Wing et al., 1989), has one XbaI site but no XhoI site.

(C) DNA gel blot analysis with the *ER* cDNA as a probe revealed a genomic rearrangement in a fast-neutron allele, *er-105*. One microgram of total genomic DNA was digested with XbaI (Xb) and ScaI (Sc). In **(B)** and **(C)**, molecular length standards are indicated at left in kilobases.

residue at subdomain VIa, into lysine (Figure 6C). Nine of 10 previously reported plant RLKs have isoleucine at this position (Walker, 1994). One exception, TMK1, has leucine, which is similar to isoleucine (Figure 6C) (Chang et al., 1992). Therefore, it is most likely that L*er* disrupts proper activity of a functional receptor kinase.

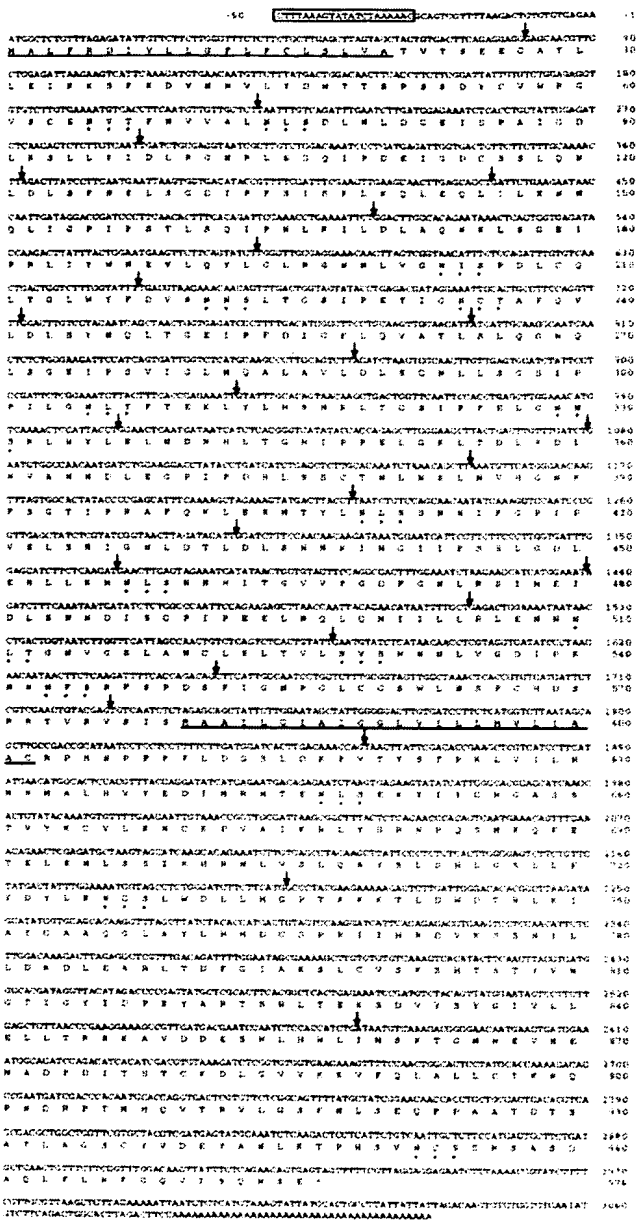


Figure 5. Nucleotide and Deduced Amino Acid Sequence of the *ER* Gene.

The positions of introns, determined by comparison of cDNA and genomic DNA sequences, are indicated by arrows. The putative signal peptide and membrane-spanning regions are underlined. The possible N-glycosylation sites (Asn-X-Ser/Thr) are indicated by asterisks. The open box indicates nucleotides that are missing in the genomic DNA of *er-104* due to a T-DNA insertion. The first nucleotide of the translation initiation site and the first amino acid of the ER protein are numbered 1. The GenBank/EMBL/DDBJ accession numbers for the wild-type *ER* cDNA and genomic DNA sequences are U47029 and D83257, respectively.

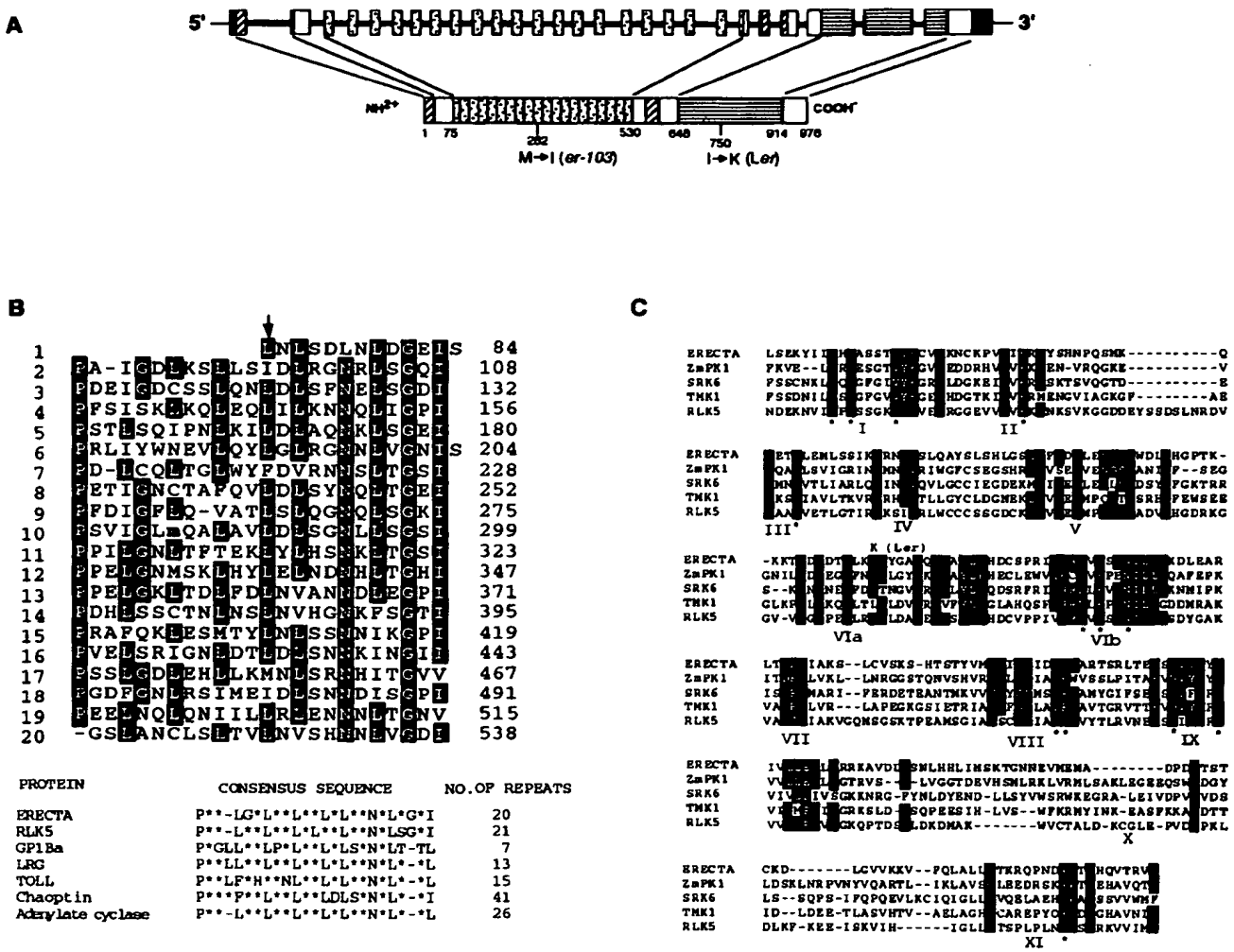


Figure 6. Structure of the Predicted ER Protein and Conserved Features of ER Amino Acid Sequence.

(A) Structure of the *ER* gene and the predicted ER protein. The boxes correspond to the exons, and the thick lines correspond to the introns. The hatched boxes represent a signal sequence and a transmembrane region, the dotted boxes represent LRRs, the horizontally lined boxes represent the putative protein kinase domain, and the closed boxes represent untranslated regions. Numbers at the bottom indicate the starting and ending positions of amino acids in each domain. The amino acids mutated in two different alleles are also indicated.

(B) The alignment of LRR repeats in the ER protein. Residues that appear at each position at >50% frequency are shown by black boxes. The positions of introns are indicated by an arrow. An amino acid mutated in *er-103* is indicated by a boldface lowercase m. At the bottom is a comparison of the LRR consensus sequence of ER with the consensus sequences of other LRR-containing proteins (Kobe and Deisenhofer, 1994). Asterisks designate nonconserved amino acids, and dashes designate gaps.

(C) Homology alignments of the cytoplasmic kinase domain of ER with the kinase domains of ZmPK1 (Walker and Zhang, 1990), SRK6 (Stein et al., 1991), TMK1 (Chang et al., 1992), and RLK5 (Walker, 1993), the four receptor-like protein kinases in plants. Highly conserved residues identical to those in ER (more than three of four) are indicated by black boxes. The positions of 11 protein kinase subdomains (Hanks and Quinn, 1991) are indicated by Roman numerals, and the invariant amino acids (Hanks and Quinn, 1991) are indicated by asterisks. Dashes designate gaps. The mutation observed in *Ler* is indicated above the altered residue, amino acid 750.

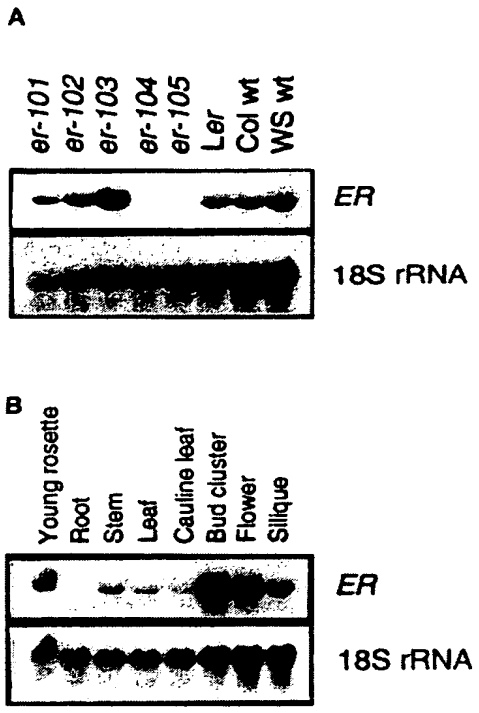


Figure 7. Expression of *ER*.

(A) Expression of *ER* transcripts of different mutant alleles. Total RNA was isolated from inflorescences of 5-week-old wild-type (wt) plants (Col and WS), *Ler* plants, and plants homozygous for each *er* allele. Five micrograms of total RNA was loaded in each lane. The membrane was probed with the *ER* cDNA that hybridized with a 3.3-kb band, and then the same blot was reprobed with the 18S rDNA as a control.
(B) Expression of *ER* transcripts of different organs. Wild-type Col plants (3 to 5 weeks old) were dissected before RNA extraction. "Young rosette" designates the aerial parts of 2-week-old seedlings. Five micrograms of total RNA was loaded in each lane. The blot was probed with the *ER* cDNA and reprobed with the 18S rDNA as a control.

The *ER* Gene Is Most Highly Expressed at or around the Shoot Meristem

The expression of *ER* transcripts was analyzed for different *er* alleles and in the major plant organs by RNA gel blot analysis (Figures 7A and 7B). As previously described, no detectable transcript of any size was observed in *er-104* and *er-105* (Figure 7A). In *er-104*, T-DNA was inserted at the 5′ untranslated region by deleting 28 nucleotides, which possibly include a transcriptional initiation point. Therefore, the gene may not be transcribed. In *er-105*, ~4 kb of DNA of unknown origin was inserted within the *ER* locus (Figures 4A and 4C); consequently, the transcript may be unstable. A transcript of normal size was detected in RNA gel blots from the other four *er* alleles, including two point mutation alleles *er-103* and *Ler* (Figure 7A). This result is consistent with the DNA gel blot analysis in which those four alleles displayed the exact same pattern as those of wild-type Col and WS (data not shown).

ER was expressed at the highest level in young floral buds, including inflorescence meristems (i.e., bud clusters), at high levels in flowers and siliques, and at lower levels in stems, rosette leaves, and cauline leaves; no expression was observed in roots (Figure 7B). This expression pattern is consistent with the phenotype conferred by *er*. *ER* transcripts are more abundant in younger leaves than in mature rosette leaves (Figure 7B), suggesting that *ER* may be involved in the cell expansion process during leaf formation.

DISCUSSION

All of the five *er* alleles that we have isolated from Col and WS ecotypes showed phenotypes similar to that conferred by *Ler*, although a precise analysis showed the different degrees of severity (Figures 1 and 2). *er* plants have very short inflorescence stems, siliques, and pedicels (Figure 1). However, it is unlikely that *ER* participates in the general process of cell elongation for the following reasons. First, the *er* mutation affects not only elongation of the organs. Regarding inflorescence development, the *er* mutation seems to alter the timing of stem elongation and flower bud formation; as a consequence, flowers and flower buds are clustered at the top of the inflorescence (Figure 2). Such an alteration in inflorescence architecture is not observed in cell elongation mutants, such as *dwarf* and *auxin resistant2* (Bowman, 1993). Second, the defects of the *er* mutation are restricted to the above-ground portion of the plants and are most conspicuous in inflorescence stems and siliques. Third, *ER* transcripts are not uniformly expressed throughout the plants. The expression is highest at or around the apical meristem and is absent in roots (Figure 7B). These observations led us to presume that ER plays a role in coordination of cell growth patterns within the organ primordia initiated from the shoot apical meristem.

The predicted structure of ER supports the hypothesis that ER participates in the coordination of cell growth patterns. The *ER* gene encodes a putative transmembrane receptor protein kinase with extracellular ligand binding domain (Figures 6A to 6C). This strongly suggests that ER participates in intercellular signal transduction, perhaps through interaction with extracellular ligands that activate the intracellular kinase domain. In plants, several transmembrane RLKs have been identified (reviewed in Walker, 1994). Based on the structural similarity of the extracellular domains, they are classified into three groups, namely, the S domain, epidermal growth factor–like domain, and LRR domain (reviewed in Walker, 1994). None of these domains has been shown to function as a receptor, nor have their ligands been identified. The S domain was originally found in self-incompatibility locus glycoproteins (SLGs) in Brassica. The physical linkage of S domain RLKs

and SLGs in Brassica suggests that they function as a pair in the self-incompatibility recognition between pollen and stigma (Stein et al., 1991). TMK1 and RLK5 have been reported as receptor-like serine/threonine protein kinases of Arabidopsis with extracellular LRR domains, although their biological functions are unknown (Chang et al., 1992; Walker, 1993).

LRR motifs are found in a variety of proteins with diverse functions and are thought to be involved in protein–protein interactions (Kobe and Deisenhofer, 1994). Extracellular LRRs are found in numerous transmembrane and membrane-attached proteins. Functionally, they are classified into two groups. One group contains several families of signal transduction receptors. The mammalian luteinizing hormone/chorionic gonadotropin (LH/CG) receptor and follicle-stimulating hormone (FSH) receptor participate in transmembrane signal transduction of peptide hormones (McFarland et al., 1989; Heckert et al., 1992). A series of experiments, such as the truncation of the LH/CG receptor and domain swapping between the LH/CG receptor and the FSH receptor, has identified the extracellular LRR domain as a specific binding site for peptidic hormones (Braun et al., 1991). The human Trk tyrosine kinase receptor is encoded by a proto-oncogene that has both an extracellular LRR domain and a cytoplasmic kinase domain (Schneider and Schweiger, 1991). In plants, the tomato disease resistance gene *Cf-9* encodes a membrane-anchored glycoprotein with an extracellular LRR motif, which is most likely the receptor domain for the Avr9 peptide, the fungus avirulence gene product (Jones et al., 1994).

The second group contains adhesive proteins. Human GPIb, a receptor for the von Willebrand factor, mediates the adhesion of platelets (Lopez et al., 1987, 1988). The role of the LRR domain in adhesion is suggested because the polypeptide fragment of GPIb containing LRR binds to the von Willebrand factor (Handa et al., 1986; Titani et al., 1987). Drosophila Toll, chaoptin, and connectin have extracellular LRRs, and their functions are required to orient cells during development (Hashimoto et al., 1988; Reinke et al., 1988; van Vactor et al., 1988; Nose et al., 1992). The ectopic expression of Toll, chaoptin, and connectin in Drosophila culture cells causes cell aggregation (Keith and Gay, 1990; Krantz and Zipursky, 1990; Nose et al., 1992). These experiments have demonstrated that they are indeed cell adhesion molecules. Toll protein also has a signaling role in the establishment of dorsal–ventral polarity in the Drosophila embryo via the interaction with a soluble extracellular ligand, a processed form of the Spätzle protein (Schneider et al., 1994). Thus, there may be LRR proteins that have both signaling and adhesive functions.

Considering the structural analogies of ER to mammalian receptors and cell adhesion molecules, it is not unreasonable to speculate that ER has a function similar to that of its animal counterparts. The first possibility is that ER mediates a signal from the apical meristem, perhaps by binding to a hypothetical peptidic ligand, which is analogous to animal growth factors. Alternatively, *ER* may directly promote cell membrane–cell wall attachment and regulate cell shapes. The adhesion of animal cells is mediated by the interaction of plasma membrane receptors, collectively named integrins, and a family of extracellular matrix glycoproteins that include the von Willebrand factor, fibronectins, and vitronectins (Schindler et al., 1989). A previous study using soybean cells suggested that an analogous system is involved in cell membrane–cell wall attachment in plants (Schindler et al., 1989). It is intriguing that ER shares a similar LRR motif with a receptor for the von Willebrand factor (Lopez et al., 1987, 1988).

Whatever the case may be, a cytoplasmic kinase domain in ER suggests the presence of a target molecule phosphorylated by ER. Protein phosphatase (designated KAPP, a kinase-associated type 2C protein phosphatase) was identified by interaction cloning with RLK5. RLK5 and KAPP associate in a phosphorylation-dependent manner (Stone et al., 1994). It is possible that ER also transduces signals via association with protein phosphatase, because ER shares high similarity with RLK5 in both the extracellular LRR and cytoplasmic kinase domain. The highest expression of *ER* transcripts in bud clusters (Figure 7B) suggests that the major site of interaction of ER with an unidentified ligand is at or around the shoot meristem, although we still do not know about the post-transcriptional regulation of *ER*. This interaction may regulate the subsequent development of organs, such as leaves derived from a vegetative meristem, stems derived from an inflorescence meristem, and siliques derived from a floral meristem.

The gene structure of *ER* provides clues to the evolution of the LRR domain. In the sequences encoding all 20 LRRs of *ER*, introns are present at the exact same position, between the second and third nucleotide coding leucine at position 13 (underlined) for the consensus P**LG*L**L**L*L**N*L*G*I (asterisks represent nonconserved amino acids) (Figures 5 and 6B). Interestingly, eight of the 10 LRRs of the mammalian LH/GC and the FSH receptors and two of 10 LRRs of the sea anemone G protein–coupled receptor are disrupted by the introns at the homologous position, again between the second and third nucleotide for the leucine at position 14 (underlined) of their consensus P**AF**L**L**L*L**N*L**I (Heckert et al., 1992; Tsai-Morris et al., 1992; Nothacker and Grimmelikhuijzen, 1993). The consistent and repeated location of introns in *ER* and in genes encoding animal receptors has implications for *LRR evolution. First, the LRR motif may have evolved by exon* duplication. Second, a single LRR unit may be the ancestral form for both the animal and plant kingdoms. The presence of an intron in a single LRR-encoding sequence could provide an opportunity to truncate the LRR domain, possibly by alternative splicing, to change the specificity for the particular ligand. The alternative splicing has been observed in the *rat LH/CG receptor and the sea anemone receptor (Koo et al.*, 1991; Aatsinki et al., 1992; Nothacker and Grimmelikhuijzen, 1993). To date, there is no evidence for alternative splicing of *ER* transcripts.

In this study, we have shown that *ER* encodes a putative transmembrane receptor kinase that regulates organ shape. Defining components of the signal transduction pathway mediated by ER is our next major challenge. Currently, we are

performing a genetic screen to isolate additional mutations that show a genetic interaction with *er*. These genetic approaches together with the complementary molecular approach, such as a search for molecules that interact with the LRR and the protein kinase domain of ER, may elucidate the nature of cell–cell communication controlling the development of plants.

METHODS

Isolation of *erecta* Alleles

er-101 was isolated from x-ray–irradiated seed populations of *Arabidopsis thaliana* ecotype Colombia (Col). *er-102* and *er-105* were isolated from fast-neutron–irradiated Col seed populations (Lehle Seeds, Round Rock, TX). *er-103* was isolated from ethylmethane sulfonate–treated Col seed mutagenized in our laboratory. *er-104* was isolated from a T-DNA–mutagenized population of ecotype Wassilewskija (WS) as follows. T. Oosumi and R.F. Whittier established a collection of ~500 T-DNA–transformed Arabidopsis lines in the WS ecotype by *Agrobacterium tumefaciens*–mediated plant transformation based on the method of Chang et al. (1994). The transformation vector was pGDW32 and is described by Wing et al. (1989). Screening was performed based on the phenotype of inflorescence. Mutant alleles were backcrossed twice into wild-type Col before the experiments described here. Genetic analysis confirmed a single recessive nature of the mutations that are allelic to Landsberg *erecta* (*Ler*) (data not shown). Plants were grown under continuous light.

Scanning Electron Microscopy

Samples were fixed in FAA (50% ethanol, 3.7% formaldehyde, and 5% acetic acid) overnight at 4°C and dehydrated through a graded ethanol series. Samples were critical point dried with liquid CO_2, mounted, sputter coated with gold, and viewed in a scanning electron microscope (model HCP-2; Hitachi, Tokyo, Japan).

Isolation of the *ER* Gene

Plant DNA was isolated by using the cetyltrimethylammonium bromide method (Watson and Thompson, 1986), and yeast DNA was isolated as described by Ausubel et al. (1989). An Arabidopsis Col λYES cDNA library (a gift of J. Mulligan and R. Davis, Department of Biochemistry, Stanford University, Stanford, CA) was screened to isolate the *ER* gene. Initially, 375,000 plaques were screened by using HindIII fragments of the rescued plasmids (6Ea and 6Xc) as probes. Five clones corresponding to the 0.6-kb transcripts and one partial clone (2.4 kb) corresponding to the 3.3-kb transcripts were isolated. In vivo excision of DNA was performed according to Elledge et al. (1991). Another 375,000 plaques were screened by using a partial *ER* clone (EcoRI digests of pKUT100; insert size of 2.4 kb) as a probe. Five additional clones were isolated, and one of them (3.2 kb) encoded a single open reading frame. To isolate wild-type genomic clones, a P1 library (Liu et al., 1995) was screened by using the EcoRI-XbaI fragments of the rescued plasmids 6Ea and 6XC as probes (designated L and R, respectively; see Figure 4A). One clone (61H10) hybridized with both probes, and another (28D7) hybridized only with the R probe. The 3.8-kb region of overlap between these two clones included the 0.6-kb transcript coding region as well as the 5′ end of the 3.3-kb transcript coding region.

Isolated cDNAs and two XbaI fragments of P1 genomic clone 61H10 were subcloned into pBluescriptII SK+ (Stratagene) for sequencing. Sequencing was performed using a Taq Dye Primer Cycle sequencing kit (Applied Biosystems, Foster City, CA) on an automated DNA sequencer (model 373A; Applied Biosystems). Sequencing of cDNA and genome P1 clones revealed a nucleotide base substitution within the *ER* coding region, which does not affect the amino acid sequence (data not shown). This is likely due to the difference of the population of plants used to construct the cDNA and genome P1 libraries, even though they are both designated as ecotype Col. For DNA gel blot analysis, plant and yeast genomic DNAs were digested with appropriate restriction enzymes, separated on 0.7% agarose gels, and alkali-blotted onto Hybond N+ membrane (Amersham, Arlington Heights, IL). DNA gel blot hybridizations under stringent conditions were performed as recommended by the manufacturer (Amersham). DNA probes were random prime-labeled with phosphorus-32 dCTP by a BcaBEST labeling kit (TaKaRa, Kyoto, Japan).

Isolation of the *er* Gene from *Ler* and Sequencing of *er* Alleles

The Arabidopsis *Ler* λZAPII cDNA library (a gift of K. Goto, Institute of Chemical Research, Kyoto University, Kyoto, Japan) was used to isolate the mutant *er* gene. Approximately 250,000 plaques were screened, and two partial clones (1.9 and 1.7 kb) were isolated. Purified phage were excised in vivo, according to the manufacturer's instruction (Stratagene), and generated plasmids (pKUT170 and pKUT180) were subcloned and sequenced. Alternatively, total RNA was isolated from aerial parts of *Ler* and *er-103*, and cDNA was synthesized using a first-strand cDNA synthesis kit (Amersham). First-strand cDNA was used as a template for the polymerase chain reaction to amplify the *er* cDNA for 40 cycles at 93°C for 1 min, 54°C for 2 min, and 70°C for 1.5 min with Takara Taq (TaKaRa). Sequencing was performed with the Taq DyeDeoxy Terminator Cycle sequencing kit (Applied Biosystems).

RNA Gel Blot Analysis

Total RNA was prepared by a general SDS–phenol method, with modifications as follows. After LiCl precipitation, RNA samples were dissolved in TE (10 mM Tris, 1 mM EDTA, pH 8.0) containing 0.3 M sodium acetate, stored at 0°C for 25 min, and centrifuged. Isopropanol (0.5 volume) was added to the supernatants, and total RNA was obtained as a precipitate. These steps effectively eliminated polysaccharides, especially in samples from roots and siliques. RNA was denatured by glyoxal; electrophoresis, transfer of RNA to nylon membranes (GeneScreen Plus; Du Pont), and hybridization under stringent conditions were performed as recommended by the manufacturer. *ER* cDNA (EcoRI fragment of pKUT160) was random prime-labeled with phosphorus-32 by a BcaBEST labeling kit and used as a probe. An insert from a clone containing pea 18S rDNA sequences (Jorgenson et al., 1987) was used as a control.

ACKNOWLEDGMENTS

We thank Kimiko Shinozaki for technical assistance with scanning electron microscopy; Drs. John Mulligan, Ronald Davis, and Koji Goto for gifts of cDNA libraries; Dr. Ichiro Masai for technical advice; and Dr. Nobuaki Hayashida for help with the sequence data base search. We also are grateful to Drs. Vivian Irish, Caren Chang, Albrecht von Arnim, Chris Day, and Tim McNellis for helpful comments on the manuscript and to Dr. David Jones for valuable discussion about the LRR motif. The results of experiments shown in Figures 1, 2D and 2E, 4C, 6C, and 7 (control hybridizations) were done in Dr. Xing-Wang Deng's laboratory at the Department of Biology, Yale University. K.U.T. is indebted to Dr. Deng and members of the Deng laboratory for their understanding and warm encouragement during preparation of the manuscript. K.U.T. was the recipient of a fellowship from the Japan Society for the Promotion of Science.

Received December 7, 1995; accepted February 2, 1996.

REFERENCES

Aatsinki, J.T., Pietila, E.M., Lakkakorpi, J.T., and Rajaniemi, H.-J. (1992). Expression of the LH-GC receptor gene in rat ovarian tissue is regulated by an extensive alternate splicing of the primary transcript. Mol. Cell. Endocrinol. **84**, 127–135.

Anderson, M., and Mulligan, B. (1992). *Arabidopsis* mutant collection. In Methods in Arabidopsis Research, C. Koncz, N.-H. Chua, and J. Schell, eds (Singapore: World Scientific), pp. 419–437.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K., eds (1989). Current Protocols in Molecular Biology. (New York: Greene Publishing Associates/John Wiley and Sons).

Becraft, P.W., and Freeling, M. (1994). Genetic analysis of *Rough sheath1* developmental mutants of maize. Genetics **136**, 295–311.

Bowman, J., ed (1993). Arabidopsis: An Atlas of Morphology and Development. (New York: Springer-Verlag).

Braun, T., Schofield, P.R., and Sprengel, R. (1991). Amino-terminal leucine-rich repeats in gonadotropin receptors determine hormone specificity. EMBO J. **10**, 1885–1890.

Carpenter, R., and Coen, E.S. (1995). Transposon induced chimeras show that *floricaula*, a meristem identity gene, acts non-autonomously between cell layers. Development **121**, 19–26.

Chang, C., Schaller, G.E., Patterson, S.E., Kwok, S.F., Meyerowitz, E.M., and Bleecker, A.B. (1992). The TMK1 gene from Arabidopsis codes for a protein with structural and biochemical characteristics of a receptor protein kinase. Plant Cell **4**, 1263–1271.

Chang, S.S., Park, S.K., Kim, B.C., Kang, B.J., Kim, D.U., and Nam, H.G. (1994). Stable genetic transformation of *Arabidopsis thaliana* by *Agrobacterium* inoculation *in planta*. Plant J. **5**, 551–558.

Dudley, M., and Poethig, R.S. (1993). The heterochronic *Teopod1* and *Teopod2* mutations of maize are expressed non–cell-autonomously. Genetics **133**, 389–399.

Elledge, S.J., Mulligan, J.T., Ramer, S.W., Spottswood, M., and Davis, R. (1991). λYes: A multifunctional complementary DNA expression vector for the isolation of genes by complementation of yeast and *Escherichia coli* mutations. Proc. Natl. Acad. Sci. USA **88**, 1731–1735.

Handa, M., Titani, K., Holland, L.Z., Robert, J.R., and Ruggeri, Z.M. (1986). The von Willebrand factor-binding domain of platelet membrane glycoprotein Ib. J. Biol. Chem. **261**, 12579–12585.

Hanks, S.K., and Quinn, A.M. (1991). Protein kinase catalytic domain sequence database: Identification of conserved features of primary structure and classification of family members. Methods Enzymol. **200**, 38–62.

Hantke, S.S., Carpenter, R., and Coen, E.S. (1995). Expression of *floricaula* in single cell layers of periclinal chimeras activates downstream homeotic genes in all layers of floral meristems. Development **121**, 27–35.

Hashimoto, C., Hudson, K.L., and Anderson, K.V. (1988). The Toll gene of Drosophila, required for dorsal–ventral embryonic polarity, appears to encode a transmembrane protein. Cell **52**, 269–279.

Heckert, L.L., Daley, I.J., and Griswold, M.D. (1992). Structural organization of the follicle-stimulating hormone receptor gene. Mol. Endocrinol. **6**, 70–80.

Horn, M.A., and Walker, J.C. (1994). Biochemical properties of the autophosphorylation of RLK5, a receptor-like protein kinase from *Arabidopsis thaliana*. Biochem. Biophys. Acta **1208**, 65–74.

Hwang, I., and Goodman, H.M. (1995). An *Arabidopsis thaliana* root-specific kinase homolog is induced by dehydration, ABA, and NaCl. Plant J. **8**, 37–43.

Hwang, I., Kohchi, T., Hauge, B.M., Goodman, H.M., Schmidt, R., Cnops, G., Dean, C., Gibson, S., Iba, K., Lemieux, B., Arondel, V., Danhoff, L., and Somerville, C. (1991). Identification and map position of YAC clones comprising one-third of the Arabidopsis genome. Plant J. **1**, 367–374.

Jackson, D., Veit, B., and Hake, S. (1994). Expression of maize *Knotted1* related homeobox gene in the shoot apical meristem predicts patterns of morphogenesis in the vegetative shoot. Development **120**, 405–413.

Jones, D.A., Thomas, C.M., Hammond-Kosack, K.E., Balint-Kurti, P.J., and Jones, J.D.G. (1994). Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvim* by transposon tagging. Science **266**, 789–793.

Jorgenson, R.A., Cuellar, R.E., Thompson, W.F., and Kavanagh, T.A. (1987). Structure and variation in ribosomal RNA genes of pea: Characterization of a cloned rDNA repeat and chromosomal rDNA variants. Plant Mol. Biol. **8**, 3–12.

Keith, F.J., and Gay, N. (1990). The *Drosophila* membrane receptor Toll can function to promote cellular adhesion. EMBO J. **9**, 4299–4306.

Kobe, B., and Deisenhofer, J. (1994). The leucine-rich repeats: A versatile binding motif. Trends Biochem. Sci. **19**, 415–420.

Komeda, Y., and Torii, K.U. (1994). Genetic control of stem formation. In Molecular Mechanisms Determining the Plant Form, A. Watanabe, H. Fukuda, K. Shimamoto, and S. Naito, eds (Tokyo: Shujun-Sha Publishing), pp. 34–40 (in Japanese).

Koo, Y.B., Ji, I., Slaughter, R.G., and Ji, T.H. (1991). Structure of the luteinizing hormone receptor gene and multiple exons of the coding sequence. Endocrinology **128**, 2297–2308.

Krantz, D.E., and Zipursky, S.L. (1990). Drosophila chaoptin, a member of the leucine-rich repeat family, is a photoreceptor cell-specific adhesion molecule. EMBO J. **9**, 1969–1977.

Liu, Y.-G., Mitsukawa, N., Vazquez-Tello, A., and Whittier, R.F. (1995). Generation of a high-quality P1 library of *Arabidopsis* suitable for chromosome walking. Plant J. **7**, 351–358.

Lopez, J.A., Chung, D.W., Fujikawa, K., Hagan, F.S., Papayannopoulou, T., and Roth, G.J. (1987). Cloning of the α chain of human platelet glycoprotein Ib: A transmembrane protein with homology to a leucine-rich glycoprotein. Proc. Natl. Acad. Sci. USA **84**, 5615–5619.

Lopez, J.A., Chung, D.W., Fujikawa, K., Hagan, F.S., and Roth, G.J. (1988). The α and β chains of human platelet glycoprotein Ib are both transmembrane proteins containing a leucine-rich amino acid sequence. Proc. Natl. Acad. Sci. USA **85**, 2135–2139.

Ma, H., Yanofsky, M.F., and Meyerowitz, E.M. (1990). Molecular cloning and characterization of *GPA1*, a G protein α subunit gene from *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **87**, 3821–3825.

McFarland, K.C., Sprengel, R., Phillips, H.S., Köhler, M., Rosemblit, N., Nikolics, K., Segaloff, D.L., and Seeburg, P.H. (1989). Lutropin-choriogonadotropin receptor: An unusual member of the G protein–coupled receptor family. Science **245**, 494–499.

Nose, A., Mahajan, V.B., and Goodman, C.S. (1992). Connectin: A homophilic cell adhesion molecule expressed on a subset of muscle and the motorneurons that innervate them in Drosophila. Cell **70**, 553–567.

Nothacker, H.P., and Grimmelikhuijzen, C.J.P. (1993). Molecular cloning of a novel, putative G-protein coupled receptor from sea anemone structurally related to members of the FSH, TSH, LH-GC receptor from mammals. Biochem. Biophys. Res. Commun. **197**, 1062–1069.

Rédei, J.P. (1992). A note on Columbia wild type and Landsberg *erecta*. In Methods in Arabidopsis Research, C. Koncz, N.-H. Chua, and J. Schell, eds (Singapore: World Scientific), p. 3.

Reinke, R., Krantz, D.E., Yen, D., and Zipursky, S.L. (1988). Chaoptin, a cell surface glycoprotein required for Drosophila photoreceptor cell morphogenesis, contains a repeat motif found in yeast and humans. Cell **52**, 291–301.

Schindler, M., Meiners, S., and Cheresh, D.A. (1989). RGD-dependent linkage between plant cell wall and plasma membrane: Consequences for growth. J. Cell Biol. **108**, 1955–1965.

Schneider, D.S., Jin, Y., Morisato, D., and Anderson, K.V. (1994). A processed form of the Spätzle protein defines dorsal–ventral polarity in the *Drosophila* embryo. Development **120**, 1243–1250.

Schneider, R., and Schweiger, M. (1991). A novel modular mosaic of cell adhesion motifs in the extracellular domains of the neurogenic *trk* and *trkB* tyrosine kinase receptors. Oncogene **6**, 1807–1811.

Sinha, N., and Hake, S. (1990). Mutant characters of *Knotted* maize leaves are determined in the innermost tissue layers. Dev. Biol. **141**, 203–210.

Stein, J.C., and Nasrallah, J.B. (1993). A plant receptor-like gene, the S-locus receptor kinase of *Brassica oleracea* L., encodes a functional serine/threonine kinase. Plant Physiol. **101**, 1103–1106.

Stein, J.C., Howlett, B., Boyes, D.C., Nasrallah, M.E., and Nasrallah, J.B. (1991). Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of *Brassica oleracea*. Proc. Natl. Acad. Sci. USA **88**, 8816–8820.

Stone, J.M., Collinge, M.A., Smith, R.D., Horn, M.A., and Walker, J.C. (1994). Interaction of a protein phosphatase with an *Arabidopsis* serine-threonine protein kinase. Science **266**, 793–795.

Szymkowiak, E.J., and Sussex, I.M. (1992). The internal meristem layer (L3) determines floral meristem size and carpel number in tomato periclinal chimeras. Plant Cell **4**, 1089–1100.

Titani, K., Takio, K., Handa, M., and Ruggeri, Z.M. (1987). Amino acid sequence of the von Willebrand factor-binding domain of platelet membrane glycoprotein Ib. Proc. Natl. Acad. Sci. USA **84**, 5610–5614.

Tsai-Morris, C.H., Buczko, E., Wang, W., Xie, X.-Z., and Dufau, M.L. (1991). Structural organization of the rat luteinizing hormone LH receptor gene. J. Biol. Chem. **266**, 11355–11369.

Tsukaya, H., Naito, S., Rédei, J.P., and Komeda, Y. (1993). A new class of mutation in *Arabidopsis thaliana*, *acaulis1*, affecting the development of both inflorescence and leaves. Development **118**, 751–764.

van Vactor, D., Jr., Krantz, D.E., Reinke, R., and Zipursky, S.L. (1988). Analysis of mutants in chaoptin, a photoreceptor cell-specific glycoprotein in Drosophila, reveals its role in cellular morphogenesis. Cell **52**, 281–290.

von Heijne, G. (1983). Patterns of amino acid near signal-sequence cleavage sites. Eur. J. Biochem. **133**, 17–21.

Walker, J.C. (1993). Receptor-like protein kinase genes of *Arabidopsis thaliana*. Plant J. **3**, 451–456.

Walker, J.C. (1994). Structure and function of the receptor-like protein kinases of higher plants. Plant Mol. Biol. **26**, 1599–1609.

Walker, J.C., and Zhang, R. (1990). Relationship of a putative receptor protein kinase from maize to the S-locus glycoprotein of *Brassica*. Nature **345**, 743–746.

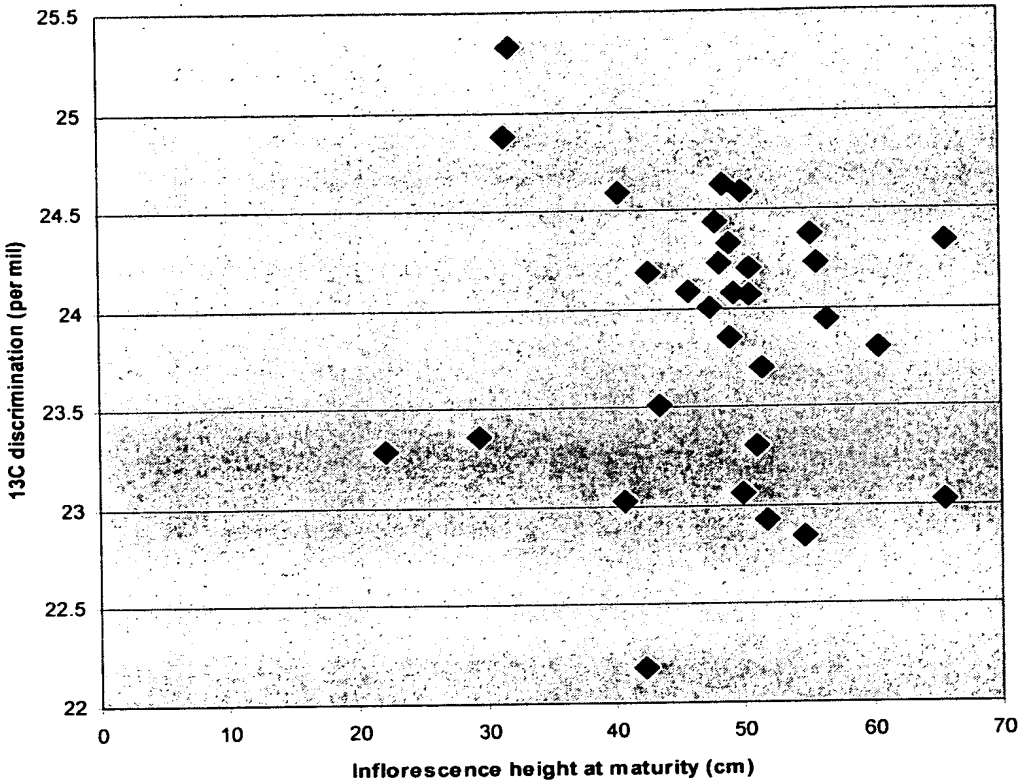
Watson, J.C., and Thompson, W.F. (1986). Purification and restriction endonuclease analysis of plant nuclear DNA. Methods Enzymol. **118**, 57–75.

Weinstein, J.N., Blumenthal, R., van Renswoude, J., Kempf, C., and Klausner, R.D. (1982). Charge clusters and the orientation of membrane proteins. J. Membr. Biol. **66**, 203–212.

Wing, D., Koncz, C., and Schell, J. (1989). Conserved function in *Nicotiana tabacum* of a single Drosophila Hsp70 promoter heat shock element when fused to a minimal T-DNA promoter. Mol. Gen. Genet. **219**, 9–16.

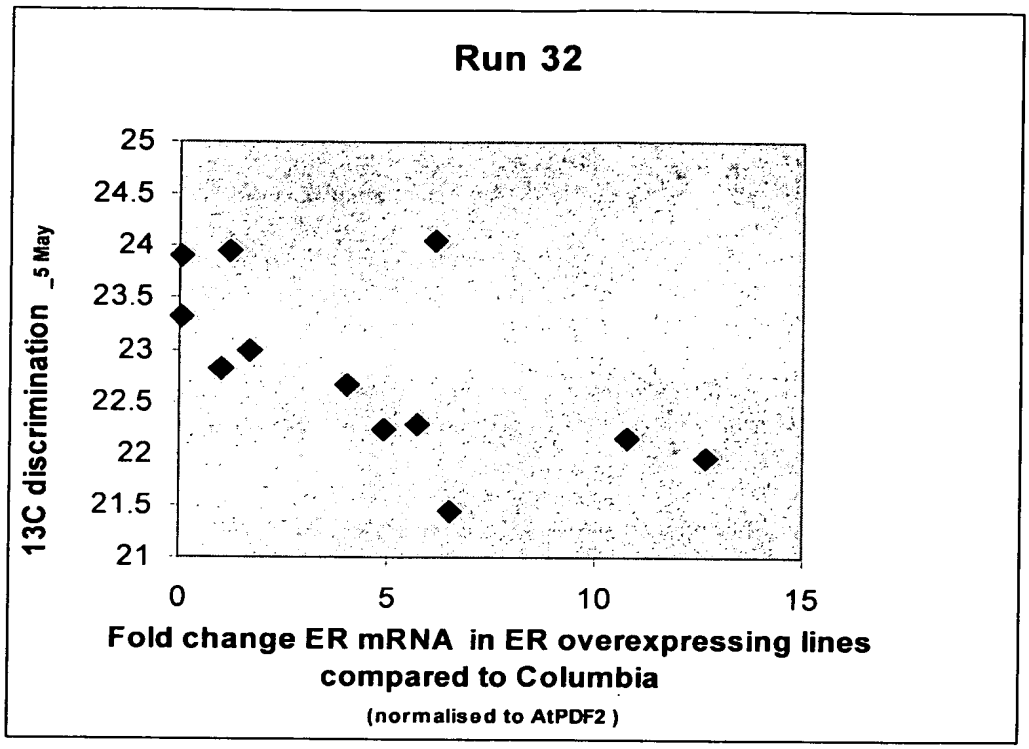
EXHIBIT F

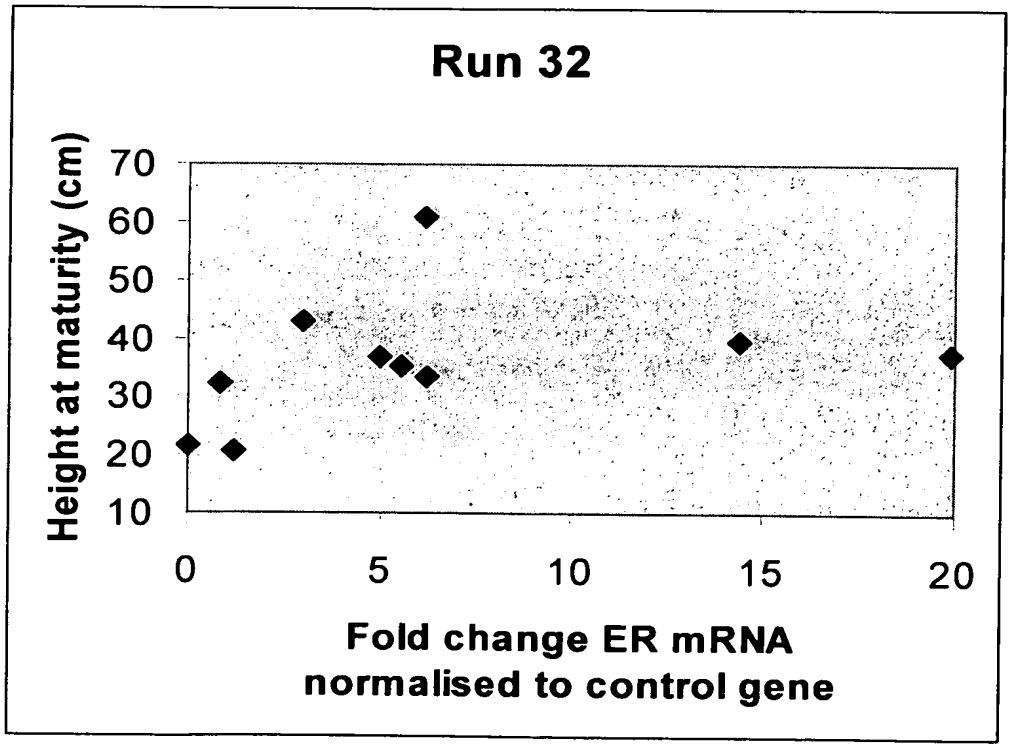
Run 35- range of Arabidopsis accessions polymorphic in *ERECTA*



Applicant: Josette Masle et al.
Serial No.: 10/519,135
Filed: August 15, 2005
Exhibit F

EXHIBIT G


Run 32
13C discrimination _5 May
25
24.5
24
23.5
23
22.5
22
21.5
21
0
5
10
15
Fold change ER mRNA in ER overexpressing lines compared to Columbia
(normalised to AtPDF2)


Run 32
Height at maturity (cm)
70
60
50
40
30
20
10
0
5
10
15
20
Fold change ER mRNA
normalised to control gene